

METHOD FOR THE TRANSFORMATION OF VEGETABLE PLASTIDS**Publication number:** AU2002366865**Publication date:** 2003-07-09**Inventor:** BIESGEN CHRISTIAN**Applicant:** SUNGENE GMBH AND CO KGAA**Classification:****- international:** **C07K14/18; C12N15/82; C07K14/005; C12N15/82;**
(IPC1-7): C12N15/82; C12N5/14**- European:** C07K14/18F; C12N15/82A12**Application number:** AU20020366865 20021216**Priority number(s):** DE20011063159 20011220; WO2002EP14303
20021216**Also published as:**

WO03054201 (A1)

EP1461439 (A1)

EP1461439 (A0)

Report a data error here

Abstract not available for AU2002366865

Abstract of corresponding document: **WO03054201**

The invention relates to new methods for producing transgenic plants with genetically modified plastids and the transgenic plants produced by said methods.

Data supplied from the **esp@cenet** database - Worldwide

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES
PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum
Internationales Büro



(43) Internationales Veröffentlichungsdatum
3. Juli 2003 (03.07.2003)

PCT

(10) Internationale Veröffentlichungsnummer
WO 03/054201 A1

(51) Internationale Patentklassifikation⁷: **C12N 15/82**,
5/14

(21) Internationales Aktenzeichen: PCT/EP02/14303

(22) Internationales Anmeldedatum:
16. Dezember 2002 (16.12.2002)

(25) Einreichungssprache: Deutsch

(26) Veröffentlichungssprache: Deutsch

(30) Angaben zur Priorität:
101 63 159.6 20. Dezember 2001 (20.12.2001) DE

(71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme von
US): **SUNGENE GMBH & CO. KGAA** [DE/DE]; Cor-
rensstrasse 3, 06466 Gatersleben (DE).

(72) Erfinder; und

(75) Erfinder/Anmelder (nur für US): **BIESGEN, Christian**
[DE/DE]; Reichenstr. 29, 06484 Quedlinburg (DE).

(74) Anwalt: **BIEBERBACH, Andreas**; BASF Aktienge-
sellschaft, 67056 Ludwigshafen (DE).

(81) Bestimmungsstaaten (national): AE, AG, AL, AM, AT,
AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR,
CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE,
GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR,
KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK,
MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU,
SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA,
UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Bestimmungsstaaten (regional): ARIPO-Patent (GH,
GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),
eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ,
TM), europäisches Patent (AT, BE, BG, CH, CY, CZ, DE,
DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT,
SE, SI, SK, TR), OAPI-Patent (BF, BJ, CF, CG, CI, CM,
GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Veröffentlicht:

— mit internationalem Recherchenbericht

*Zur Erklärung der Zweibuchstaben-Codes und der anderen
Abkürzungen wird auf die Erklärungen ("Guidance Notes on
Codes and Abbreviations") am Anfang jeder regulären Ausgabe
der PCT-Gazette verwiesen.*

WO 03/054201 A1

(54) Title: METHOD FOR THE TRANSFORMATION OF VEGETABLE PLASTIDS

(54) Bezeichnung: VERFAHREN ZUR TRANSFORMATION VON PFLANZLICHEN PLASTIDEN

(57) Abstract: The invention relates to new methods for producing transgenic plants with genetically modified plastids and the transgenic plants produced by said methods.

(57) Zusammenfassung: Die vorliegende Erfindung betrifft neue Verfahren zur Herstellung transgener Pflanzen mit gentechnisch veränderten Plastiden, sowie die mit diesen Verfahren hergestellten transgenen Pflanzen.

Verfahren zur Transformation von pflanzlichen Plastiden

Beschreibung

5

Die vorliegende Erfindung betrifft neue Verfahren zur Herstellung transgener Pflanzen mit gentechnisch veränderten Plastiden, sowie die mit diesen Verfahren hergestellten transgenen Pflanzen.

- 10 Ziel biotechnologischer Arbeiten an Pflanzen ist die Herstellung von Pflanzen mit vorteilhaften, neuen Eigenschaften zum Beispiel zur Steigerung der landwirtschaftlichen Produktivität, zur Qualitätssteigerung bei Nahrungsmitteln oder zur Produktion bestimmter Chemikalien oder Pharmazeutika.

15

Plastiden stellen Organellen innerhalb pflanzlicher Zellen mit einem eigenen Genom dar. Sie haben eine wesentliche Funktion in der Photosynthese sowie der Aminosäure- und Lipidbiosynthese. Das Genom der Plastiden besteht aus einer doppelsträngigen, zirkulären DNA mit einer durchschnittlichen Größe von 120 bis 160 kb

- 20 und liegt - z.B. in Blattzellen - mit ca. 1900 bis 50000 Kopien pro Zelle vor (Palmer (1985) Ann Rev Genet 19:325-54). Ein einzelnes Plastid trägt dabei eine Kopienzahl von ca. 50 bis 100. Der Begriff der Plastiden umfasst Chloroplasten, Proplastiden,

- 25 Etioplasten, Chromoplasten, Amyloplasten, Leukoplasten und Elaioplasten (Heifetz P (2000) Biochimie 82:655-666). Die verschiedenen Formen sind in einander überführbar und entstehen alle aus den Proplastiden. Daher enthalten alle Ausprägungsformen der Plastiden die gleiche Erbinformation. Bevorzugt werden in der

- 30 Literatur jedoch grüne Zellen, die Chloroplasten als Ausprägungsform enthalten, als Ausgangsmaterial für die Transformation von Plastiden genutzt.

Neben Tabak wurde die Plastidentransformation in Kartoffel

- 35 (Sidorov VA et al. (1999) Plant J 19:209-216; WO 00/28014), Petunie (WO 00/28014), Reis (Khan MS und Maliga P (1999) Nature Biotech 17:910-915; WO 00/07431; US 6,153,813), Arabidopsis (Sikdar SR et al. (1998) Plant Cell Reports 18: 20-24; WO 97/32977) und Raps (WO 00/39313) gezeigt (Übersichtsartikel: 40 Bogorad L (2000) TIBTECH 18:257-263). Kürzlich wurden auch transplastome Tomatenpflanzen beschrieben (Ruf S et al. (2001) Nature Biotech 19:870- 875).

Es ist für die Pflanzenbiotechnologie von großem wirtschaftlichen

- 45 Interesse, effiziente Methoden für die Plastidentransformation zu entwickeln (McFadden G (2001) Plant Physiol. 125:50-53). Die

2

stabile Transformation von Plastiden höherer Pflanzen gehört dabei zu den großen Herausforderungen.

Die bei der Insertion in die nukleäre DNA oft angewandte

- 5 Technik der ungezielten (illegitimen) DNA-Insertion hat bei der Plastidentransformation den Nachteil, dass mit sehr hoher Wahrscheinlichkeit auf dem gendichten plastidären Genom ein essentielles Gen getroffen wird, was häufig letal für die Pflanze wäre. In Plastiden ist daher eine gezielte Insertion von Fremd-
10 DNA vorteilhaft.

Verschiedene Verfahren zur gezielten Insertion in das plastidäre Genom sind beschrieben. Zuerst wurde die Plastidentransformation in Grünalgen beschrieben (Boynton JE et al. (1988) Science 240:

- 15 1534-1538; Blowers AD et al. (1989) Plant Cell 1:123-132), später in höheren Pflanzen wie Tabak (Svab Z et al. (1990) Proc Natl Acad Sci USA 87:8526-8530).

EP-A 0 251 654, US 5,932,479, US 5,451,513, US 5,877,402,

- 20 WO 01/64024, WO 00/20611, WO 01/42441, WO 99/10513, WO 97/32977, WO 00/28014, WO 00/39313 beschreiben Methoden und DNA-Konstrukte zur Transformation von Plastiden höherer Pflanzen, wobei die zu transformierende DNA über homologe Rekombination ("Doppel-Crossover") in das Plastom (Plastidengenom) eingebaut wird. Im
25 allgemeinen werden auf jeder Seite der zu insertierenden Sequenz homologe Regionen von 1000 bp oder mehr genutzt. Dies führt sehr schnell zu großen Vektoren, deren Handhabung wenig komfortabel ist. Darüber hinaus sinkt die Effizienz der Transformation. Mit zunehmender Länge der zu integrierenden Fremd-DNA sinkt die
30 Effizienz der homologen Rekombination. Ein weiterer Nachteil ist die Tatsache, dass für jede Pflanzenart ein homologer Bereich identifiziert werden muss, der für den Prozess der DNA Integration mittels Doppel-Crossover genutzt werden kann. WO 99/10513 beansprucht, eine intergenische DNA Sequenz identifiziert zu
35 haben, die ausreichend homolog zwischen den Genomen der Chloroplasten vieler höherer Pflanzen sei und so als universelle Zielsequenz fungieren kann. Dass dieser Vektor bei anderen Arten als Tabak erfolgreich genutzt werden kann, wurde indies nicht nachgewiesen. Im Gegenteil: In WO 01/64024 passt derselbe Erfinder
40 den Transformationsvektor an Nicht-Tabak Pflanzenarten an, indem er aus diesen isolierte homologe DNA-Sequenzen verwendet. Da bei allen der oben beschriebenen Verfahren nur wenige Rekombinationsereignisse resultieren, ist eine Selektion der rekombinanten plastidären DNA-Moleküle erforderlich.

3

- Die Plastiden-DNA höherer Pflanzen liegt in bis zu mehreren tausend Kopien pro Zelle vor. Um eine stabile Integration der Fremd-DNA zu erreichen, müssen alle Kopien der Plastiden-DNA gleichermaßen verändert werden. Im Fall der Plastidentrans-
- 5 formation spricht man davon, den homotransplastomen Zustand erreicht zu haben. Diesen Zustand erreicht man durch einen sogenannten "segregation and sorting" Prozess, indem man die Pflanzen unter ständigen Selektionsdruck hält. Durch den anhaltenden Selektionsdruck werden während der Zell- und Plastidenteilung
- 10 solche Plastiden angereichert, in denen bereits viele Kopien der plastidären DNA verändert wurden. Der Selektionsdruck wird aufrecht gehalten, bis der homotransplastome Zustand erreicht ist (Guda C et al. (2000) Plant Cell Reports 19:257-262). Es ist eine große Herausforderung, alle Kopien des Plastidengenoms zu ver-
- 15 ändern, um homotransplastome Pflanzen zu erhalten, die auch über Generationen ohne Zugabe eines Selektionsmittels das Fremdgen stabil in ihr plastidäres Genom eingebaut haben (Bogorad L (2000) TIBTECH 18:257-263). Zusätzlich zu dem ständigen Selektionsdruck wird gegebenenfalls durch wiederholte Regeneration bereits trans-
- 20 formierten Gewebes das Erreichen des homotransplastomen Zustand sichergestellt (Svab Z und Maliga P (1993) Proc Natl Acad Sci USA 90:913-917). Dies jedoch schränkt das Pflanzenmaterial, welches für die Plastidentransformation zur Verfügung steht, ein. Es wird daher vorgeschlagen, das Transgen ggf. mit einem anderen Gen,
- 25 welches essentiell für das Überleben der Pflanze ist, zu koppeln. Allgemein wird als einzige Pflanzenart, deren Plastiden effizient transformiert werden können, Tabak angesehen. Die Limitierung der Technik wird vor allem in der gegenwärtig zur Verfügung stehenden Gewebekultur und den Regenerationsprotokollen gesehen (Ruf, S.
- 30 et al. 2001 Nature Biotech. 19: 870-875). Gewebekultur-Techniken und Selektionsprozesse sind meist nicht universell auf alle Pflanzenarten anwendbar und stellen eine wesentliche Limitierung der Plastidentransformation dar, die insbesondere die Übertragbarkeit der Methode auf andere Arten als Tabak betrifft (Kota M
- 35 et al. (1999) Proc Natl Acad Sci USA 96:1840- 1845). Eine kürzlich veröffentlichte Transformation von Plastiden von Tomate basiert auf Modifikationen im Regenerations- und Selektionsschema (Ruf S et al. (2001) Nature Biotech 19:870-875), die jedoch kosten- und zeitintensiv sind. Ein anderer Ansatz zielt darauf
- 40 ab, die Anzahl der Plastiden pro Zelle und der DNA Moleküle pro Plastid zu reduzieren, um so weniger DNA Moleküle modifizieren zu müssen (Bogorad L (2000) TIBTECH 18:257-263). Insgesamt sind die Selektions- und Segregationsprozesse sehr zeitaufwendig.
- 45 WO 99/10513 beschreibt ein Verfahren bei dem auf dem zu transformierenden Plasmid ein plastidärer ORI (origin of replication, Replikationsursprung) lokalisiert wird, um so die Anzahl der für

die Integration in das Plastidengenom zur Verfügung stehenden Kopien des zu transformierenden Vektors zu erhöhen (Guda C et al. (2000) Plant Cell Reports 19:257-262).

- 5 Die Notwendigkeit die Technik der Plastidentransformation zu verbessern wird auch bei Heifetz und Tuttle erwähnt (Heifetz P und Tuttle AM (2001) Curr Opin Plant Biol 4:157-161). WO 00/32799 lehrt, die Effizienz der Plastidentransformation durch den Einsatz von Pflanzen mit vergrößerten Plastiden zu erhöhen. Dadurch
- 10 ergibt sich eine große Oberfläche der Plastiden, durch die die zu transformierende DNA leichter in die Plastiden eindringen kann. Der Mechanismus der DNA Integration erfolgt aber auch hier mittels konventioneller homologer Rekombination wie bei den oben beschriebenen Verfahren.
- 15 Die homologe Rekombination selber wird nicht als Ansatzpunkt für weitere Optimierungsmöglichkeiten diskutiert, da sie in Plastiden als hoch und damit nicht limitierend angesehen wird (Bock R und Hagemann R (2000) Progress in Botany. 61:76-90; Maliga P et al. (1994) Homologous recombination and integration of foreign DNA in plastids of higher plants. In: Homologous recombination and gene silencing in plants. Paszkowski J ed., Kluwer Academic publishers, S.83-93; Carrer H und Maliga P (1995) Bio/Technology 13: 791-794).
- 25 Allgemein werden also optimierte Gewebe- und Regenerations-techniken als Ansatz favorisiert, die jedoch den Nachteil haben, dass sie nicht universell auf alle Pflanzenarten anwendbar sind, sondern jeweils an deren Eigenarten angepasst werden müssen. Nur
- 30 zwei Arbeiten schlagen vor, auch durch eine Verbesserung der Insertionshäufigkeit der zu transformierenden DNA in die DNA von Organellen die Transformation dieser zu verbessern. In dem einen Fall wird dabei keine technische Lehre beschrieben, wie man eine Verbesserung der Integrationshäufigkeit in plastidäre DNA erreichen kann (van Bel AJE et al. (2001) Curr Opin Biotechnol. 12: 144-149). Odom OW et al. (2001 Mol. Cell Biol. 21: 3472-3481) schlagen allgemein vor, die Transformationseffizienz durch Verwendung selten schneidende Restriktionsendonuklease zu steigern.
- 40 Sequenzspezifische Rekombinasen bewirken - in Abhängigkeit von der Lokalisation und Orientierung der entsprechenden Erkennungssequenzen - eine Insertion, Excision, Translokation oder Inversion von Nukleinsäurefragmenten. Es wird zwischen den Familien der Integrasen, wie beispielsweise Cre oder Flp, und
- 45 der Familie der Resolvasen/Invertasen, wie beispielsweise die

ΦC31 Rekombinase, die TP901-1 Rekombinase, die Gin Rekombinase, die Hin Rekombinase etc. unterschieden.

Diese Enzyme erkennen in der Regel relativ kurze, spezifische Nukleinsäuresequenzen, die sowohl der Erkennung als auch der Rekombination dienen. Beispielfhaft seien die Cre Rekombinase (Sternberg und Hamilton (1981) J Mol Biol 150:467-486), die Flp Rekombinase (Brosch et al. (1982) Cell 29:227-234) und die R Rekombinase (Matsuzaki et al. (1990) J Bacteriology 172:610-618) zu nennen.

Die Cre Rekombinase erlaubt es dem Bacteriophagen P1, während seines Lebenszyklus ein monomeres Phagengenom zu erhalten, indem die während der Replikation erzeugten Sequenzduplikationen durch die Rekombinase ausgeschnitten werden (Guo et al. (1997) Nature 389:40-46). Cre ist ein 38 kD Protein und katalysiert die Rekombination zwischen zwei 34 Basenpaarsequenzen (loxP).

loxP 5'-ATAACTTCGTATA GCATACAT TATACGAAGTTAT-3'

LoxP besteht aus zwei 13 Basenpaar palindromischen Sequenzen getrennt von einer 8 Basenpaar Kernsequenz ("Spacer"). Die "Repeat"-Sequenzen fungieren als Bindesequenzen der Cre-Rekombinase, wohingegen der Sequenzschnitt ("Kreuzung der zu rekombinierenden Sequenzen") in der "Spacer" Region erfolgt. Vermutlich wird ein Übergangskomplex aus vier Cre-Molekülen und zwei loxP Sequenzen gebildet (Guo et al. (1997) Nature 389:40-46). Die Asymmetrie der "Spacer" Region bedingt die Richtung der Rekombination. Sind die beiden Rekombinationssequenzen gleichgerichtet, so erfolgt Integration bzw. Excision. Ist die Orientierung gegenläufig, so erfolgt Inversion (Yang and Mizuuchi (1997) Structure 5:1401-1406). Cre ist in einer Vielzahl von Organismen aktiv, einschließlich Pflanzen (Albert et al. (1995) Plant J 7:649-659; Dale und OW (1990) Gene 91:79-85; Odell et al. (1990) Mol Gen Genet 223:369-378). Cre kann eine sequenzspezifische Insertion beispielsweise in Hefe oder Säugerzellen bewirken (Sauer und Henderson (1990) New Biol 2:441-449).

Das Flp Inversionssystem der Hefe *S. cerevisiae* (2μ Plasmid) basiert auf der Nutzung einer Int-Flp Rekombinase. Flp erkennt eine Region - genannt FRT - bestehend zwei jeweils 13 Basenpaar langen invertierten "Repeat"-Sequenzen, getrennt durch eine Sequenz (Spacer) von 8 bp. Ein in der natürlichen Sequenz weiterhin vorliegender dritter Repeat scheint für die Funktionalität nicht essentiell zu sein. Essentiell ist eine invertierte Lokalisation der Repeats #1 und #2. Der versetzte endonukleolytische Schnitt erfolgt innerhalb der 8 bp Sequenz. Die Länge des Spacers

6

ist wichtig. Eine Insertion von weiteren 10 bp kann jedoch toleriert werden. Das System basiert auf zwei flankierenden Zielsequenzen.

5 #1 spacer #2 #3
 FRT 5-GAAGTTCCTATaC TTTCTAGA GAATAGGAACTTC [C GAATAGGAACTTC]-3'

Wie Cre ist auch Flp in verschiedenen Systemen funktionell, einschließlich im pflanzlichen Zellkern (Lyznik et al. (1993)

10 Nucl Acids Res 21:969-975).

Natürlicherweise sind viele Rekombinationssysteme reversibel. Beim Nutzen der Rekombinasen wie Cre oder FLP liegt das Gleichgewicht der Reaktion weit auf der Seite der Excision. Daher sind
 15 diese Systeme - ohne weitere Modifikation - nicht für die Integration von Fremd-DNA, sondern nur für die Excision (z.B. Marker-excision) geeignet. So wurde das Cre-lox System des Phagen P1 genutzt, um den Selektionsmarker aadA wieder aus dem plastidären Genom zu entfernen (Hajdukiewicz PTJ et al. (2001) Plant J
 20 27:161-170; Corneille S et al. (2001) Plant J 27: 171-178; WO01/21768; WO01/29241). Es gibt bisher keine Beschreibungen zum Nutzen von Rekombinasen bei intermolekularen Reaktionen, insbesondere zur Integration von Fremd-DNA, in Plastiden. Wirtsorganismus-unabhängige Rekombinationssysteme sind in Plastiden
 25 bislang lediglich im Zusammenhang einer Deletion definierter Sequenzen aus dem Plastom beschrieben worden.

Die Nutzung von sequenzspezifischen Rekombinasesystemen zur Integration heterologer Sequenzen in die chromosomale DNA (u.a.
 30 von Pflanzen) ist beschrieben. Dabei handelt es sich - im Unterschied zu der Excision, die intramolekular verläuft - um ein intermolekulares Ereignis. So wurde versucht, die Integrationsrate im Vergleich zur Excisionsrate zu erhöhen, indem bei einem Rekombinationsereignis zwischen zwei nicht identischen aber noch
 35 von der Rekombinase genutzten Erkennungsregionen wenigstens eine hybride Rekombinase-Erkennungsstelle entsteht, die nicht mehr oder nur sehr ineffizient von der entsprechenden Rekombinase genutzt werden kann. Damit wird die Excision des gerade integrierten DNA Fragmentes unterdrückt (Sauer B (1994) Curr Opin
 40 Biotechnol 5:521-527) und eine Stabilisierung der Integration erreicht (Albert H et al. (1995) Plant J 7:649-659). Ähnliches wird durch eine gezielte, zeitlich eng begrenzte Expression der entsprechenden Rekombinase - wie Cre oder Flp - erreicht (Metzger D und Feil R (1999) Curr Opin Biotechnol 10:470-476; Albert H
 45 et al. (1995) Plant J 7:649-659).

Die aktuelle Forschung zur Optimierung der Plastidentransformation richtet sich nicht auf eine Optimierung der homologen Rekombination, sondern zum Beispiel auf verbesserte Selektionsmarker, verbesserte Selektions- und Regenerationstechniken etc.

5

- Die Rekombinase des Phagen Φ C31 wurde schon für biotechnologische Applikationen in der Vergangenheit eingesetzt. US 5,190,871 beschreibt ein Verfahren, mit dessen Hilfe ein Plasmid mit Hilfe der Erkennungsstellen-spezifischen Integrationsfunktion des
- 10 Phagen Φ C31 in das Genom eines Streptomyceten eingebaut wird. Die Integration erfolgt an die in diesen Wirtszellen natürlich vorhandene Erkennungsregion attB. WO 00/11155 nutzt Rekombinasen zum Einbau von Fremd-DNA in eukaryotische Genome. Dabei werden die Rekombinasen unter der Kontrolle eines eukaryotischen Promotors
- 15 exprimiert. Die Rekombination erfolgt dabei zwischen einer Erkennungsregion auf dem zu transformierenden Plasmid und pseudo-Erkennungsregionen, die sich nativ im Genom des Eukaryoten befinden. In WO 00/11155 wird eine Erkennungsregion attB oder attP in das Genom einer eukaryotischen Zelle eingebracht. Dies wird
- 20 nur für tierische Zellen gezeigt und es gibt keine technische Offenbarung, dass und wie man solche Erkennungsregionen auch in Organellen einbringen könnte. Im Gegensatz zum Kerngenom, welches in der angesprochenen Erfindung betrachtet wurde, indem nur eine (heterozygot) bis zwei (homozygot) Genomkopien vorliegen, gibt
- 25 es viele Plastiden pro Zelle und viele DNA-Kopien pro Plastid.

- Während WO 00/11155 nur Untersuchungen in Bakterien und tierischen Zellen durchführt, beschreibt WO 01/07572 das Nutzen solcher Rekombinasen auch im Kerngenom von Pflanzen. Auch hier
- 30 wird nicht gezeigt, dass solche Rekombinasen die DNA Struktur der Plastiden-DNA erkennen und dort eine effiziente Anwendung aufgrund der Vielzahl und räumlichen Trennung dieser DNA Moleküle möglich ist. Tatsächlich stellte sich eine solche Aufgabe dem Erfinder auch gar nicht, da er ein Verfahren suchte, das eine
- 35 Alternative für die im Kerngenom nur mit sehr geringen Frequenzen auftretende Homologe Rekombination darstellt.

- Unterschiedliche Rekombinasen bevorzugen unterschiedliche Konformationen der DNA, die sie als Substrate nutzen (Sadowski
- 40 PD (1986) J Bacteriol 165:341-347). Es ist daher unklar ob und welche Rekombinasen - neben dem Cre-lox System - in Plastiden funktionieren. Es wurden bisher noch nie Rekombinasen der Resolvase/Invertase Familie in Plastiden höherer Pflanzen eingesetzt. Da die Nukleosomen gänzlich anders aufgebaut sind als die
- 45 Chromosomen im Zellkern, konnte auch nicht aus der erfolgreichen Verwendung von bestimmten Rekombinasen im pflanzlichen Zellkern

davon ausgegangen werden, dass diese auch mit plastidärer DNA als ein Substrat genutzt werden können.

Die sogenannte Erkennungsstellen-spezifische Integration (*site-specific integration*) wurde bisher nur im Kerngenom von Pflanzen und Tieren angewendet. Ziel der im Stand der Technik beschriebenen Arbeiten war die Bereitstellung eines Alternativverfahrens zu der homologen Rekombination für eine sequenzspezifische Integration in das Kerngenom. Im Kerngenom höherer Eukaryonten sind Systeme wie Cre oder FLP allenfalls so effizient wie die dort normalerweise auftretende illegitime Rekombination, meist jedoch wesentlich ineffizienter (Sauer B (1994) Curr Opin Biotechnol 5:521-527). Im Unterschied zu Kerngenom liegt das Plastom in zahlreichen, meist mehreren tausend Kopien pro Zelle vor, so dass hier besonders hohe Anforderungen an die Effizienz der Insertion gestellt werden müssen.

Im Unterschied zu Plastiden ist die Insertion in die Kern-DNA mittels homologer Rekombination problematisch und erfolgt meist mittels zufälliger illegitimen Integration. Dies zeigt, dass im Kern-Genom etablierte Techniken nicht unbedingt auf die Plastiden übertragen werden können. Bei Plastiden höherer Pflanzen findet im Gegensatz zu der Situation im Kern nahezu ausschließlich und mit hoher Effizienz eine Integration über homologe Rekombination statt (Bock R und Hagemann R (2000) Progress in Botany 61:76-90; Maliga P et al. (1994) Homologous recombination and integration of foreign DNA in plastids of higher plants. In Homologous recombination and gene silencing in plants. Paszkowski J, ed. (Kluwer Academic publishers), S. 83-93).

Die Effizienz der homologen Rekombination zur DNA-Integration in das Plastom wurde allgemein nicht als limitierender Faktor angesehen und - im Gegenteil - als unkritisch eingestuft, so dass der Einsatz von Rekombinasen zur Integration von Fremd-DNA in die Plastiden-DNA überflüssig erscheint.

Wie die oben beschriebenen Verfahren und Probleme bei der Plastidentransformation deutlich hervorheben, ist die Bereitstellung neuer Verfahren zur Herstellung homotransplastomer Pflanzen ein lange bestehendes ungelöstes Bedürfnis der Pflanzenbiotechnologie. Ein weiteres Bedürfnis ist die Vermeidung von Antibiotika- oder Herbizidselektionsmarkern aus Gründen der Zulassung und Verbraucherakzeptanz. Für die Plastidentransformation ist bisher noch kein Verfahren beschrieben, dass ohne einen solchen Selektionsmarker auskommt.

Es ergab sich daher die Aufgabe, neue Verfahren zu entwickeln, die eine effiziente Integration von Fremd-DNA in alle Kopien der Plastiden-DNA sicher stellt und so eine effiziente Selektion entsprechender homotransplastomer Pflanzen ermöglicht. Darüber hinaus stellt sich die Aufgabe, effiziente Verfahren bereitzustellen für die Integration insbesondere großer DNA Fragmente in das Plastom in einem einzigen Transformationsschritt. Diese Aufgabe wurde durch Bereitstellung des erfindungsgemäßen Integrationsverfahrens gelöst.

10

Ein erster Gegenstand der Erfindung betrifft ein Verfahren zur sequenzspezifischen Integration einer DNA-Sequenz in die plastidäre DNA eines pflanzlichen Organismus oder von diesem abgeleitete Zellen, dadurch gekennzeichnet, dass

15

a) die plastidären DNA-Moleküle besagten pflanzlichen Organismus oder von diesem abgeleitete Zelle mindestens eine Rekombinationssequenz R1 enthalten und

20 b)

i) mindestens ein Transformationskonstrukt umfassend eine Insertionssequenz zumindest einseitig flankiert von mindestens einer weiteren Rekombinationssequenzen R2 sowie

25

ii) mindestens eine sequenzspezifische Rekombinase geeignet zur Induktion von Rekombinationseignissen an besagten Rekombinationssequenzen R1 und R2

30

in mindestens einem Plastid besagten pflanzlichen Organismus oder von diesem abgeleiteten Zellen zusammengebracht werden, und

c) die Insertionssequenz in die plastidäre DNA unter Rekombination von R1 mit R2 inseriert, und

35

d) Pflanzen oder Zellen isoliert werden, bei denen die Insertionssequenz in die plastidären DNA-Moleküle inseriert wurde.

40 Das System ermöglicht überraschenderweise eine erhebliche Steigerung der Effizienz bei der Herstellung überwiegend homotransplastomer Pflanzen. Dabei wird sowohl die Effizienz der Insertion in die plastidäre DNA als auch die Effizienz des Selektionsprozess überwiegend homotransplastomer Pflanzen gesteigert.

45 Besonders vorteilhaft ist, dass bei Ausnutzen des erfindungsgemäßen Verfahrens auf homologe Bereiche in den Transformationsvektoren verzichtet werden kann. Dadurch werden die Vektoren

10

kleiner und leichter handhabbar. Ferner wird die Effizienz der Plastidentransformation und damit die Gesamteffizienz des Verfahrens gesteigert. Letzteres ist insbesondere dann von Interesse, wenn im Rahmen der Insertionssequenz große DNA-Fragmente in das Plastom eingebracht werden sollen, was - beispielsweise - im Fall der Einbringung mehrerer Gene gegeben ist. Dadurch ist das erfindungsgemäße Verfahren, der im Stand der Technik beschriebenen Insertion mittels homologer Rekombination (z.B. Doppel-Crossover) überlegen, da beim letzteren die Effizienz der Transformation (und somit die Gesamteffizienz) durch die zusätzlich erforderlichen Homologiebereiche reduziert wird. Somit löst die vorliegende Erfindung das Problem der Integration großer DNA Fragmente (> 4kb) in einem einzigen Transformationsschritt mit ausreichend hoher Effizienz.

15

Vorteile des hier beschriebenen Verfahrens sind: Es muss nur einmal eine sogenannte Masterpflanze - d.h. eine Pflanze, die in Ihrem Plastom eine Rekombinationssequenz R1 enthält und als Ausgangspflanze für das erfindungsgemäße Verfahren fungieren kann - in einer entsprechenden Art erstellt werden. Es können auch nicht-modifizierte Pflanzen als Masterpflanzen verwendet werden, soweit deren Plastom natürlicherweise eine Erkennungssequenz für eine Rekombinase enthält. Die Plastiden der vorgenannten Masterpflanzen können anschließend mit Hilfe des erfindungsgemäßen, wirtsunabhängigen Integrationsmechanismus transformiert werden. Dadurch kann man die Effizienz der Plastidentransformation in den Pflanzenspezies steigern, in denen die Integration der Fremd-DNA in das Plastom ein limitierender Faktor ist. Auf den Transformationsvektoren werden nur wenige Basenpaar lange Sequenzabschnitte benötigt, die den Integrationsort in das Plastom bestimmen. Im Gegensatz dazu benötigt man für die Integration mittels Doppel-Crossover Sequenzabschnitte mit extensiver Homologie zum Plastom der Wirtspflanze. Damit können die Transformationsvektoren der vorliegenden Erfindung verhältnismäßig klein gehalten werden. Darüber hinaus ist die wirtsunabhängige Integration auch weitgehend unabhängig von der Länge der in die Masterpflanze einzubringenden DNA. Es können leicht sehr viele Gene oder sehr große Gene in die Organellen einer solchen Masterpflanzen mit dem hier beschriebenen Verfahren integriert werden.

40

Gegenüber der Transformation des Zellkerns hat die Transformation von Plastiden zahlreiche Vorteile. Unter anderem sind zu nennen:

- a) Während die homologe Rekombination in die nukleäre DNA nur schwer zu realisieren ist, kann in Plastiden DNA leicht an einem vordefinierten Ort mittels Doppel-Crossover, einer Form der homologer Rekombination, integriert werden. Positions-

45

11

effekte oder "Genesilencing", die man bei Kerntransformationen aufgrund der illegitimen Integration an einen nicht vordefinierten Lokus findet, werden so vermieden.

- 5 b) Es können sehr hohe Expressionslevel erreicht werden, vermutlich aufgrund der hohen Kopiezahl der plastidären DNA.
- c) Die DNA der Plastiden wird bei höheren Pflanzen in der Regel nur maternal vererbt, so dass die eingebrachte Fremd-DNA nicht über Pollen verbreitet und ein Auskreuzen somit effektiv unterbunden werden kann.
- 10
- d) Die prokaryotische Natur der Plastiden ermöglicht die Expression von Genen im Rahmen einer polycistronischen Operonstruktur. Daher ist es nicht notwendig, jedes zu exprimierende Gen mit einem eigenen Promotor etc. auszurüsten. Dies erleichtert das Einbringen vieler Gene auf einmal, etwa um ganze Biosynthesewege in die Plastiden einzubringen.
- 15
- 20 "Plastid" meint Proplastiden sowie alle daraus hervorgehenden Organellen wie beispielsweise Chloroplasten, Etioplasten, Chromoplasten, Amyloplasten, Leukoplasten, Dermaplasten und Elaioplasten (Heifetz P (2000) Biochimie 82:655-666).
- 25 "Plastom" meint das Genom, also die Gesamtheit der genetischen Information, eines Plastids.
- "Homotransplastom" meint einen transplastomen und homoplastomen
- 30 Zustand.
- "Transplastom" meint in Bezug auf beispielsweise eine Pflanze, Zelle, Gewebe, Plastid oder plastidäre-DNA alle solche durch gentechnische Methoden zustande gekommene Formen der vorgenannten,
- 35 die eine plastidäre DNA umfassen, die durch gentechnische Methoden modifiziert wurde, wobei die Modifikation beispielsweise Substitutionen, Additionen, Deletionen, Inversion oder Insertionen eines oder mehrerer Nukleotidreste umfassen kann.
- 40 "Heteroplastom" meint das Vorliegen einer gemischten Population verschiedener plastidärer DNAs innerhalb eines einzelnen Plastides oder innerhalb einer Population von Plastiden innerhalb einer pflanzlichen Zelle oder Gewebe.
- 45 "Homoplastom" meint eine einheitliche Population von plastidärer DNA innerhalb eines einzelnen Plastides oder innerhalb einer Population von Plastiden innerhalb einer pflanzlichen Zelle oder

12

Gewebe. Homoplastome Zellen, Gewebe oder Pflanzen sind genetisch stabil, da sie nur eine Art plastidärer DNA umfassen d.h. sie bleiben im allgemeinen homoplastom auch wenn den Selektionsdruck nicht mehr anhält. Durch Selbstung erhaltene Nachkommen sind
5 ebenfalls homoplastom.

Im Rahmen dieser Erfindung meint "überwiegend homoplastom" oder "überwiegend homotransplastom" all solche Pflanzen oder Zellen, bei denen der Anteil der erwünschten plastidären hinsichtlich
10 eines Merkmals veränderten DNA-Moleküle - beispielsweise mit einer Rekombinationssequenz - mindestens 50 %, bevorzugt mindestens 70 %, ganz besonders bevorzugt mindestens 90 %, am meistens bevorzugt mindestens 95 % von der Gesamtheit aller plastidären DNA-Moleküle in einer Pflanze oder einem Gewebe,
15 Zelle bzw. Plastid derselben beträgt. Überwiegend homoplastome oder überwiegend homotransplastome Pflanzen können durch weiteres Aufrechterhalten des Selektionsdruckes und gegebenenfalls wiederholenden Regenerationen in homoplastome oder homotransplastome Pflanzen umgewandelt werden. In einer besonderen Ausführungsform
20 ist daher eine überwiegend homoplastome bzw. homotransplastome Pflanze rein homoplastom bzw. homotransplastom. Eine Pflanze, die beispielsweise bezüglich einer Rekombinationssequenz überwiegend homoplastom bzw. homotransplastom oder rein homoplastom bzw. homotransplastom ist, wird infolge als "Masterpflanze" bezeichnet.
25 Der Anteil der erwünschten plastidären hinsichtlich eines Merkmals veränderten DNA-Moleküle kann beispielsweise mittels Southern Analyse - wie beispielhaft in Beispiel 4 beschrieben - in der dem Fachmann bekannten Weise ermittelt werden. Das Verhältnis zwischen den plastidären Ausgangs-DNA Molekülen
30 und den hinsichtlich eines Merkmals veränderten plastidären DNA-Molekülen lässt sich durch Vergleich der Intensität der jeweiligen Banden bestimmen.

"Pflanzlicher Organismus oder von diesem abgeleitete Zellen"
35 meint allgemein jede Zelle, Gewebe, Teile oder Vermehrungsgut (wie Samen oder Früchte) eines Organismus, der zur Photosynthese befähigt ist. Eingeschlossen sind im Rahmen der Erfindung alle Gattungen und Arten höherer und niedrigerer Pflanzen des Pflanzenreiches. Einjährige, mehrjährige, monocotyledone und
40 dicotyledone Pflanzen sind bevorzugt. Eingeschlossen sind reife Pflanze, Saatgut, Sprosse und Keimlinge, sowie davon abgeleitete Teile, Vermehrungsgut (zum Beispiel Knollen, Samen oder Früchte) und Kulturen, zum Beispiel Zell- oder Kalluskulturen. Reife Pflanzen meint Pflanzen zu jedem beliebigen Entwicklungsstadium
45 jenseits des Keimlings. Keimling meint eine junge, unreife Pflanze in einem frühen Entwicklungsstadium.

13

Bevorzugt sind Pflanzen nachfolgender Pflanzenfamilien: Amaranthaceae, Asteraceae, Brassicaceae, Carophyllaceae, Chenopodiaceae, Compositae, Cruciferae, Cucurbitaceae, Labiatae, Leguminosae, Papilionoideae, Liliaceae, Linaceae, Malvaceae, Rosaceae,

- 5 Rubiaceae, Saxifragaceae, Scrophulariaceae, Solanaceae, Sterculiaceae, Tetragoniaceae, Theaceae, Umbelliferae.

Bevorzugte monokotyle Pflanzen sind insbesondere ausgewählt aus den monokotylen Kulturpflanzen, wie zum Beispiel der Familie der

- 10 Gramineae wie Reis, Mais, Weizen oder andere Getreidearten wie Gerste, Hirse, Roggen, Triticale oder Hafer sowie dem Zuckerrohr sowie alle Arten von Gräsern.

Bevorzugte dikotyle Pflanzen sind insbesondere ausgewählt aus den

- 15 dikotylen Kulturpflanzen, wie zum Beispiel

- Asteraceae wie Sonnenblume, Tagetes oder Calendula und andere mehr,

- 20 - Compositae, besonders die Gattung Lactuca, ganz besonders die Art sativa (Salat) und andere mehr,

- Cruciferae, besonders die Gattung Brassica, ganz besonders die Arten napus (Raps), campestris (Rübe), oleracea cv Tastie

- 25 (Kohl), oleracea cv Snowball Y (Blumenkohl) und oleracea cv Emperor (Broccoli) und weitere Kohlarten; und der Gattung Arabidopsis, ganz besonders die Art thaliana sowie Kresse oder Canola und andere mehr,

- 30 - Cucurbitaceae wie Melone, Kürbis oder Zucchini und andere mehr,

- Leguminosae besonders die Gattung Glycine, ganz besonders die Art max (Sojabohne) Soja sowie Alfalfa, Erbse, Bohnengewächsen oder Erdnuss und andere mehr

35

- Rubiaceae, bevorzugt der Unterklasse Lamiidae wie beispielsweise Coffea arabica oder Coffea liberica (Kaffestrauch) und andere mehr,

- 40 - Solanaceae besonders die Gattung Lycopersicon, ganz besonders die Art esculentum (Tomate) und die Gattung Solanum, ganz besonders die Art tuberosum (Kartoffel) und melongena (Aubergine) sowie Tabak oder Paprika und andere mehr,

- 45 - Sterculiaceae, bevorzugt der Unterklasse Dilleniidae wie beispielsweise Theobroma cacao (Kakaostrauch) und andere mehr,

14

- Theaceae, bevorzugt der Unterklasse Dilleniidae wie beispielsweise *Camellia sinensis* oder *Thea sinensis* (Teestrauch) und andere mehr,
- 5 - Umbelliferae, besonders die Gattung *Daucus* (ganz besonders die Art *carota* (Karrotte)) und *Apium* (ganz besonders die Art *graveolens dulce* (Selderie)) und andere mehr; und die Gattung *Capsicum*, ganz besonders die Art *annuum* (Pfeffer) und andere mehr,
- 10 sowie Lein, Soja, Baumwolle, Hanf, Flachs, Gurke, Spinat, Möhre, Zuckerrübe und den verschiedenen Baum-, Nuss- und Weinarten, insbesondere Banane und Kiwi.
- 15 Umfasst sind ferner Schmuckpflanzen, Nutz- oder Zierbäume, Blumen, Schnittblumen, Sträucher oder Rasen. Beispielhaft aber nicht einschränkend seien zu nennen Angiospermen, Bryophyten wie zum Beispiel Hepaticae (Leberblümchen) und Musci (Moose); Pteridophyten wie Farne, Schachtelhalm und Lycopoden; Gymno-
- 20 spermen wie Koniferen, Cycaden, Ginkgo und Gnetales, die Familien der Rosaceae wie Rose, Ericaceae wie Rhododendrons und Azaleen, Euphorbiaceae wie Weihnachtssterne und Kroton, Caryophyllaceae wie Nelken, Solanaceae wie Petunien, Gesneriaceae wie das Usambaraveilchen, Balsaminaceae wie das Springkraut, Orchidaceae
- 25 wie Orchideen, Iridaceae wie Gladiolen, Iris, Freesie und Krokus, Compositae wie Ringelblume, Geraniaceae wie Geranien, Liliaceae wie der Drachenbaum, Moraceae wie Ficus, Araceae wie Philodendron und andere mehr.
- 30 Pflanzliche Organismen im Sinne der Erfindung sind weiterhin weitere photosynthetisch aktive befähigte Organismen, wie zum Beispiel Algen, Cyanobakterien sowie Moose. Bevorzugte Algen sind Grünalgen, wie beispielsweise Algen der Gattung *Haematococcus*, *Phaedactylum tricornutum*, *Volvox* oder *Dunaliella*. Insbesondere
- 35 bevorzugt ist *Synechocystis*. Cyanobakterien wie z.B. *Synechocystis* sind photosynthetisch aktive Organismen, die zwar keine Plastiden beinhalten aber quasi Plastiden darstellen und - wie die Plastiden - zahlreiche Kopien ihrer genomischen DNA enthalten. Sie gelten darüberhinaus - entwicklungsbiologisch gesehen
- 40 - als Vorläufer der heutigen Plastiden (Bauer J et al. (2001) *Cell Mol Life Sci* 58(3):420-33).
- Am meisten bevorzugt sind *Arabidopsis thaliana*, *Nicotiana tabacum*, *Tagetes* und *Brassica napus* sowie alle Gattungen und
- 45 Arten, die als Nahrungs- oder Futtermittel zum Einsatz kommen, wie die beschriebenen Getreidearten, oder sich zur Herstellung

15

von Ölen eignen, wie Ölsaaten, Nussarten, Soja, Sonnenblume, Kürbis und Erdnuss.

- "Rekombinase" meint im Rahmen dieser Erfindung sequenzspezifische
- 5 Rekombinase. Sequenzspezifische Rekombinasen umfasst Proteine, die einen reziproken Austausch von DNA Doppelsträngen zwischen zwei DNA Segmenten katalysieren können, wobei die jeweilige Rekombinase eine Präferenz für DNA Segmente mit einer bestimmten Nukleinsäuresequenz hat. Besagte miteinander rekombinierende DNA
- 10 Sequenzen sind nicht notwendigerweise identisch.

- Die im Rahmen dieser Erfindung verwendeten Rekombinasen werden insbesondere zur Integration von DNA Sequenzen verwendet. Sind für diesen Prozess - neben der Rekombinase selber - weitere
- 15 Faktoren erforderlich, so ist der Komplex aus Rekombinase und den entsprechenden Integrationsfaktoren selber auch als Rekombinase im Sinne der Erfindung zu verstehen. Beispielsweise ist eine Integration mittels der λ -Rekombinase nur dann effizient möglich, wenn auch zugleich der IHF ("integration host factor") vorhanden
- 20 ist. In einer bevorzugten Ausführungsform umfasst besagter Rekombinase-Komplex keine Faktoren, die ein Ausschneiden zuvor durch Rekombination insertierter DNA-Sequenzen begünstigen. Beispielsweise umfasst der Komplex aus λ -Rekombinase und IHF nicht noch zusätzlich den Faktor xis.

- 25 Bevorzugt ist die Geschwindigkeit einer Rekombinationsreaktion einer sequenzspezifischen Rekombinase unter Einsatz ihrer bevorzugten Erkennungssequenz (beispielsweise ihrer natürlichen Erkennungssequenz) mindestens zehnfach so hoch wie für den Ein-
- 30 satz einer beliebigen anderen nicht-homologen Sequenz, bevorzugt mindestens hundertfach so hoch, besonders bevorzugt mindestens eintausendfach so hoch, am meistens bevorzugt mindestens zehntausendfach so hoch. Dabei meint eine "nicht-homologe Sequenz" bevorzugt solche, die bei gleicher Länge, eine Sequenzidentität
- 35 zu der bevorzugten Erkennungssequenz (beispielsweise ihrer natürlichen Erkennungssequenz) der jeweiligen Rekombinase von weniger als 90 %, bevorzugt weniger als 50 %, besonders bevorzugt weniger als 30 % aufweisen.

- 40 Allgemein bevorzugt sind Rekombinasen, die alleine eine nahezu irreversible Integration katalysieren. Bevorzugt stammen diese aus Organismen, die normalerweise bei Temperaturen zwischen 10 und 45°C leben. Auch wenn die meisten bekannten Rekombinasen aus Prokaryoten, Bakteriophagen bzw. Viren isoliert wurden,
- 45 ist die Erfindung nicht auf diese beschränkt, sondern kann auch Rekombinasen aus eukaryotischen Organismen vorteilhaft einsetzen.

16

Die im Rahmen dieser Erfindung genutzte sequenzspezifische Rekombination ist nicht auf einen bestimmten Mechanismus limitiert. In der Regel erfolgt die Rekombination jedoch in zwei Schritten: Einem sequenzspezifischen Schnitt der Erkennungs-
5 sequenz gefolgt von einer Wiederverknüpfung, wobei meist ein kovalentes Protein-DNA Intermediat zwischen Zielsequenz und Rekombinase vorliegt. Für manche Rekombinationen ist allein die Rekombinase als Protein erforderlich. Andere können weitere Hilfsfaktoren benötigen. Bevorzugt ist das benutzte Rekombi-
10 nationssystem unabhängig von der Enzymausstattung des Empfängerplastids.

Besonders bevorzugt sind Rekombinasen, die zu den Familien der Integrasen oder der Resolvasen/Invertasen zählen (Stark WM et al.
15 (1992) Trends Genet 8:432-439). Integrasen (Tyrosin-Rekombinasen) haben ein Tyrosin im aktiven Zentrum, welches als Nucleophil im Spaltungsschritt fungiert und in der Bildung einer 3'-Phosphotyrosylbindung mit der DNA involviert ist. Die Reaktion erfolgt über ein "Holliday Intermediat". Resolvasen/Invertasen (Serin-
20 Rekombinasen) haben ein Serin im aktiven Zentrum, welches als Nucleophil im Spaltungsschritt fungiert und in der Bildung einer 5'-Phosphoserinylbindung mit der DNA involviert ist. Die Reaktion erfolgt nicht über ein "Holliday Intermediat".

25 Die Familie der Integrasen (auch Tyrosin-Rekombinasen genannt) umfassen beispielhaft die Cre Rekombinase aus dem Bakteriophagen P1, die FLP Rekombinase aus Hefe, die λ -Integrase des Phagen Lamda sowie die R Rekombinase des pSR1 Plasmid der Hefe *Zygosaccharomyces rouxii* (Araki et al. (1985) J Mol Biol.
30 182:191-203) und die h Rekombinase (Argos, et al. (1986) EMBO J 5:433-440, 1986; Sadowski PD (1995) Progress Nucl Acid Res Mol Biol 51:53-91).

Die λ -Integrase benötigt in der Regel weitere Faktoren, die an
35 Sequenzen benachbart zu unmittelbaren Erkennungssequenz binden. Neben der Rekombinase (auch Int für Integrase bezeichnet) wird noch IHF ("integration host factor") benötigt. Für die Umkehrreaktion, das Ausschneiden, ist zusätzlich noch der Faktor xis erforderlich.

40 Die Familie der Resolvasen/Invertasen umfasst beispielhaft die Φ C31 Rekombinase, R4 Rekombinase (GenBank Acc.-No.: D38173 Nukleotide 292 bis 1701), das Hin Inversionssystem aus *Salmonella typhimurium* und die TP901-1 Rekombinase (Hallett und Sherratt
45 (1997) FEMS Microbiol Rev 21:157-178). Weitere Familienmitglieder

umfassen die Invertasen Gin des Mu Phagen, Cin des Bakteriophagen P und Pin des ϕ 14 Phagen.

- Im Unterschied zu der Integrase-Familie haben die Mitglieder
5 der Resolvasen in der Regel eine konservierte N-terminale katalytische Domäne (Crellin und Rood (1997) J Bacteriology 15 179(16):5148-5156; Christiansen et al. (1996) J Bacteriology 178(17):5164-58173). Wie einige der Cre-typ Rekombinasen sind manche Resolvasen unabhängig von Faktoren des Wirtes (Thorpe und
10 Smith (1998) PNAS USA 95:5505-5510). Die Resolvasen/Invertasen nutzen einen anderen Reaktionsmechanismus als die Integrasen (Gopaul DN und van Duyne GD (1999) Curr Opin Structural Biol 9:14-20; Jayaram M (1994) TIBS 19:78-83). Besonders bevorzugt ist die Untergruppe der "Extended Resolvases" (Brondsted L und Hammer
15 K (1999) Appl Environm Microbiol 65: 752-758). Diese haben eine Homologie zu den anderen Resolvasen/Invertasen, weisen aber im Vergleich zu diesen eine Extension am C-Terminus auf. Der Vorteil dieser Rekombinasen liegt darin, dass sie wenig bis keine Helferfaktoren für die Katalyse der Rekombination zwischen zwei
20 Erkennungsregionen benötigen (Brondsted L und Hammer K (1999) Appl Environm Microbiol 65:752-758). Für die effiziente Katalyse der Umkehrreaktion hingegen werden weitere Faktoren benötigt. Durch Expression der entsprechenden Rekombinase alleine kann man daher bei geeigneter Positionierung der Erkennungssequenzen eine
25 unter diesen Bedingungen nahezu irreversible Integration einer heterologen DNA-Sequenz bewirken. Besonders bevorzugt aus dieser Untergruppe sind die Rekombinasen der Phagen TP901-1 und Φ C31.

- Bevorzugt werden die Rekombinasen der Integrase-Familie genutzt,
30 weil sie natürlicherweise in der Regel intermolekulare Rekombinationen vermitteln.

- Besonders bevorzugt werden in der vorliegenden Erfindung Rekombinasen eingesetzt, die zwei unterschiedliche Erkennungsregionen
35 rekombinieren und die bei der Rekombination entstehenden hybriden Erkennungsregion von der Rekombinase nicht mehr effizient oder nur mit einem weiteren Faktor erkannt und genutzt werden.

- Während bei den Rekombinationsereignissen der homologen Rekombi-
40 nation (z.B. Doppel-Crossovers) extensive Sequenzhomologie benötigt wird, vermittelt die Erkennungssequenz-spezifische Rekombination DNA-Umordnungen zwischen Segmenten, die keine extensive Homologie aufweisen. Die Rekombinationsereignisse treten bei der Erkennungssequenz-spezifischen Rekombination aus-
45 schließlich an spezifischen Stellen der DNA auf (Craig NL (1988) Annu Rev Genet 22:77-105).

- Rekombinationssysteme wie die λ Integrase aus dem λ -Phagen, die Rekombinase aus dem Mycobakteriophagen L5 (Pena CEA et al. (2000) Proc Natl Acad Sci USA 97:7760-7765; Lee MH et al. (1991) Proc Natl Acad Sci USA 88:3111-3115), die Rekombinase aus dem Phagen Φ C31 (Kuhstoss S und Rao RN (1991) J Mol Biol 222: 897-908; Groth AC et al. (2000) Proc Natl Acad Sci USA 97: 5995- 6000; Rausch H und Lehmann M (1991) Nucl Acids Res. 19: 5187- 5189; Thorpe HM und Smith MCM (1998) Proc Natl Acad Sci USA 95: 5505-5510; Kuhstoss S. et al. (1991) Gene 102:145-146; Harris JE et al. (1983) Gene 22:167-174) oder die Rekombinase aus dem Phagen TP901-1 (Brondsted L und Hammer K (1999) Appl Environm Microbiol 65:752-758; Bruener A et al. (1999) J Bacteriol 181: 7291- 7297; Koch B et al. (1997) Appl Environm Microbiol 63: 2439-2441, Christiansen B et al. (1996) J Bacteriol 178:5164-5173) nutzen natürlicherweise für ein Rekombinationsereignis zwei in ihrer Sequenz unterschiedliche Erkennungsregionen, in der Regel attP und attB genannt. Diese werden zu neuen, hybriden Erkennungsregionen (attR und attL) rekombiniert, die nicht von der entsprechenden Rekombinase allein effizient genutzt werden. Erst wenn ein oder mehrere weitere Faktoren hinzu kommen, werden diese Erkennungsregionen wieder erkannt und eine Rekombination (Rückreaktion) zwischen diesen kann erfolgen. Eine Vielzahl von Faktoren, die für die katalytische Wirkung der λ -Rekombinase eine Rolle spielen, sind bekannt (Landy A (1993) Curr Opin Gen Develop 3:699-707). Als weitere, bevorzugt verwendete Rekombinasen seien zu nennen: xisF aus Anabaena (Ramaswamy KS et al. (1997) Mol Microbiol 23:1241-1249; Carrasco CD et al. (1994) Genes Develop 8:74-83), Integrase vom Phagen Φ LC3 (Lillehaug D (1997) Gene 188:129-136), Rekombinase kodiert vom sre Gen des R4-Phagen (Matsuura M et al. (1996) J Bacteriol 178:3374-3376).

- Neben den bevorzugt genutzten Rekombinasesystemen, die allein natürlicherweise eine irreversible Reaktion vermitteln, sind auch die bekannten Rekombinase Systeme wie Cre/lox; FLP/frt aus der Hefe; R-RS aus *Zygosaccharomyces rouxii* (Onouchi, H et al., 1991, Nucl. Acids Res. 19: 6373-6378) oder Gin-gix aus dem Bakteriophagen Mu (Maeser S und Kahmann R (1991) Mol Gen Genet 230:170-176) im Rahmen dieser Erfindung nutzbar. Diese Systeme vermitteln natürlicherweise eine reversible Rekombination, wobei das Gleichgewicht meist auf die Seite der Excision verschoben ist. Für die Anwendung im Rahmen dieser Erfindung werden diese Systeme bevorzugt nahezu irreversibel gemacht (siehe unten).

- Ganz besonders bevorzugt werden die Rekombinasen aus Φ C31 und TP901-1 genutzt. Am meisten bevorzugt sind die Rekombinasen mit Nukleinsäuresequenzen hinterlegt unter den GenBank Acc.-No Y14232

19

(TP901-1; komplementär bp 30 bis 1487) bzw. X59938 (ΦC31; Nukleotide 232 bis 2073).

Es sind verschiedene Verfahren denkbar, um eine Rekombinase in
5 Plastiden einzubringen oder dort zu exprimieren. Beispielfhaft
jedoch nicht einschränkend sind zu nennen:

a) Nukleäre Expression unter Verwendung plastidärer Transit-
peptide

10

Eine Expressionskassette kodierend für eine Rekombinase kann
in der dem Fachmann bekannten Weise konstruiert, in den Zell-
kern eingeführt und - optional - stabil in die chromosomale
DNA integriert werden. Die Expression erfolgt dann transient
15 oder - bei Integration in die chromosomale DNA - stabil. Für
einen Transport in die Plastiden wird die Rekombinase bevor-
zugt in Fusion mit einer Plastidenlokalisationssequenz (PLS)
exprimiert. Verfahren, an sich nicht in den Plastiden lokali-
sierte Proteine, gezielt in die Plastiden zu transportieren
20 sowie verschiedene PLS-Sequenzen sind beschrieben (Klosgen RB
und Weil JH (1991) Mol Gen Genet 225(2):297-304; Van Breuse-
gem F et al. (1998) Plant Mol Biol 38(3):491-496). Bevorzugt
sind PLS, welche nach Translokation der Rekombinase in die
Plastiden vom Rekombinase-Teil enzymatisch abgespalten
25 werden. Insbesondere bevorzugt ist die PLS, die von der
plastidären Nicotiana tabacum Transketolase oder einem
anderen Transitpeptid (z.B. dem Transitpeptid der kleinen
Untereinheit der Rubisco oder der Ferredoxin NADP Oxido-
reduktase als auch der Isopentenylpyrophosphat Isomerase-2)
30 oder dessen funktionellem Äquivalent abgeleitet ist. Für die
nukleäre Expression sind im Prinzip alle Promotoren geeignet,
die eine Expression in Pflanzen ermöglichen. Beispiele sind
weiter unten gegeben. Bevorzugt sind konstitutive Promotoren
wie der 35S Promotor des CaMV oder der Nitrilase-1 Promotor
35 des nit1 Gens aus Arabidopsis (GenBank Acc.-No.: Y07648.2,
Nukleotide 2456 bis 4340; Hillebrand H et al. (1998) Plant
Mol Biol 36 (1):89-99; Hillebrand H et al. (1996) Gene
170(2):197-200).

40 Bevorzugte PLS Sequenzen sind:

- i) das Transitpeptid der Isopentenylisomerase (IPP) aus
Arabidopsis (GenBank Acc.-No.: NC_003074; Nukleotide
604657 - 604486)

45

20

ii) Transitpeptide abgeleitet von der kleine Untereinheit (SSU) der Ribulosebisphosphatcarboxylase (Rubisco ssu) aus beispielsweise Erbse, Mais, Sonnenblume oder Arabidopsis.

5

- Arabidopsis thaliana: GenBank Acc.-No.: z.B. AY054581, AY054552;

10

- Erbse, GenBank Acc.-No.: z.B. X00806, Nukleotide 1086 bis 1256; X04334, X04333 (Hand JM (1989) EMBO J 8(11):3195-206). Hier besonders bevorzugt: Expressionskassette und Transitpeptid (Erbse, rbcS3A) aus Vektor pJIT117 (Guerineau F (1988) Nucleic Acids Res 16(23):11380. Besonders bevorzugt ist die Peptidsequenz gemäß SEQ ID NO: 11. Am meisten bevorzugt für die Verwendung zur Konstruktion von entsprechenden Expressionskonstrukten ist die Nukleinsäuresequenz gemäß SEQ ID NO: 10.

15

20

- Mais, GenBank Acc.-No.: z.B. S42568, S42508

- Sonnenblume, GenBank Acc.-No.: Y00431, Nukleotide 301 bis 465.

25

iii) Transitpeptide abgeleitet von Genen der pflanzlichen Fettbiosynthese wie das Transitpeptid des plastidären "Acyl Carrier Protein" (ACP) (z.B. die Arabidopsis thaliana beta-ketoacyl-ACP synthetase 2; GenBank Acc.-No.: AF318307), die Stearyl-ACP-Desaturase, β -Ketoacyl-ACP Synthase oder die Acyl-ACP-Thioesterase (z.B. A.thaliana mRNA for acyl-(acyl carrier protein)thioesterase: GenBank Acc.-No.: Z36911).

30

35

iv) das Transitpeptid der GBSSI ("Starch Granule Bound Synthase I")

v) das Transitpeptid der LHCP II Gene.

40

Besonders bevorzugt ist die plastidäre Transketolase aus Tabak (SEQ ID NO: 12). Zur Expression entsprechender Fusionsproteine können im Rahmen dieser Erfindung verschiedene PLS-Nukleinsäurekassetten in den drei Leserastern als KpnI/BamHI Fragment verwendet werden (der Translationsstart (ATG-Kodon) ist in der NcoI Schnittstelle lokalisiert) (pTP09 SEQ ID NO: 13; pTP10 SEQ ID NO: 14; pTP11 SEQ ID NO: 15).

45

21

Ein weiteres Beispiel für eine vorteilhaft einzusetzende PLS ist das Transitpeptid der plastidären Isopentenylpyrophosphat Isomerase-2 (IPP-2) aus *Arabidopsis thaliana* (SEQ ID NO: 16). Ganz besonders bevorzugt können die Nukleinsäuresequenzen

5 kodierend für drei Kassetten (entsprechend den drei Leserastern) der PLS der Isopentenylpyrophosphat Isomerase-2 (IPP-2) aus *Arabidopsis thaliana* verwendet werden (EcoRV/SalI-Kassetten mit ATG; IPP-9 SEQ ID NO: 17; IPP-10 SEQ ID NO: 18; IPP-11 SEQ ID NO: 19).

10 Die erfindungsgemäßen Nukleinsäuren können synthetisch hergestellt oder natürlich gewonnen sein oder eine Mischung aus synthetischen und natürlichen Nukleinsäure-Bestandteilen enthalten, sowie aus verschiedenen heterologen Genabschnitten

15 verschiedener Organismen bestehen.

Die für das Transitpeptid kodierende Sequenz kann ganz oder teilweise die Peptidsequenz des Ursprungsproteins umfassen. Eine exakte Bestimmung der für den Transport essentiellen

20 Aminosäurereste ist nicht erforderlich, solange die Funktionalität der PLS - nämlich der Transport in die Plastiden - gewährleistet ist, und die Funktion der Rekombinase nicht gänzlich zerstört wird. Ganz besonders bevorzugt sind die nachfolgenden PLS-Sequenzen:

25 PLS1: N-MASSSSLTSLSQAILSRVPRHGSASSSQLSPSSLTFSGLKSNPNITTSTRRR TPSSAAAAAVVRSPAIRASAATETIEKTETAGS-C (SEQ ID NO: 12). Entspricht der PLS der plastidären Transketolase aus Tabak.

30 PLS2: N-MSASSLFNLPLIRLRLSLALSSSFSSFRFAHRPLSSISPRKLPNFRAFSGTAMDTKDGSRVDM-C (SEQ ID NO: 16). Entspricht der PLS der Isopentenylpyrophosphat Isomerase-2 (IPP-2), wobei das letzte Methionin bevorzugt das Start-

35 Methionin der Rekombinase ist.

Fusionsproteine aus PLS und Rekombinase sind im Rahmen dieser Erfindung unter dem Begriff der Rekombinase subsumiert. Wird eine Rekombinase nukleär exprimiert, so wird unter der

40 Rekombinase bevorzugt ein Fusionsprotein aus PLS und der Rekombinase verstanden.

b) Expression in Plastiden

45 Eine Expression in Plastiden kann auch durch direkte Einführung einer Expressionskassette für die Rekombinase in Plastiden, ggf. Integration in die plastidäre DNA und

Expression der Rekombinase erfolgen. In einer bevorzugten Ausführungsform ist diese Expressionskassette in dem die Insertionssequenz umfassenden Transformationskonstrukt enthalten. Die Expression kann aber auch von einer separaten
5 Expressionskassette aus realisiert werden.

Dabei können zum einen spezifische Plastiden- oder Chromoplasten-Promotoren - wie unten im Detail beschrieben - zum Einsatz kommen. Eine gezielte plastidäre Expression kann
10 auch erreicht werden, wenn man zum Beispiel einen viralen, bakteriellen oder Bakteriophagen Promotor verwendet, die resultierende Expressionskassette in die plastidäre DNA einbringt und die Expression dann durch die korrespondierende virale, bakterielle oder Bakteriophagen RNA-Polymerase induziert. Die korrespondierende RNA-Polymerase kann wiederum auf
15 verschiedene Art und Weise - bevorzugt durch nukleäre Transformation als Fusionsprotein mit einer PLS - in die Plastiden eingebracht werden. Entsprechende Verfahren sind beschrieben (WO 95/16783, WO 97/06250, US 5,925,806). Das Einbringen in
20 Plastiden wird bevorzugt mittels Mikroinjektion und besonders bevorzugt mittels Partikelbeschuss durchgeführt.

c) Einbringen als RNA

Die Rekombinase kann auch durch Einführung der für die Rekombinase kodierenden - beispielsweise in vitro erzeugten - mRNA z.B. über Mikroinjektion, Partikel-Bombardierung (biolistische Verfahren), Polyethylenglykol- oder Liposomen-vermittelte Transfektion in Plastiden eingebracht werden. Diese
25 Ausführungsform ist vorteilhaft, da hier keine die Rekombinase kodierenden Sequenzen im Plastom oder Genom verbleiben. Bevorzugt wird die die Rekombinase kodierende RNA durch in vitro Transkription in einer dem Fachmann bekannten Weise hergestellt.
30

35

d) Einbringen als Protein

Die Rekombinase kann direkt beispielsweise über Mikroinjektion, Partikel-Bombardierung (biolistische Verfahren),
40 Polyethylenglykol-Transfektion oder Liposomen-vermittelte Transfektion in Plastiden eingebracht werden. Diese Ausführungsform ist vorteilhaft, da hier keine die Rekombinase kodierenden Sequenzen im Plastom oder Genom verbleiben. Ein entsprechendes Verfahren ist beispielsweise beschrieben
45 bei Segal DJ et al. (1995) Proc Natl Acad Sci USA 92:806-810.

23

Die Rekombinase kann alternativ als Fusionsprotein mit dem VirE2 oder VirF Protein eines Agrobakterium sowie einer PLS in Pflanzenzellen eingeschleust werden. Entsprechende Verfahren sind beispielsweise für die Cre-Rekombinase beschrieben (Vergunst AC et al. (2000) Science 290:979-982). Diese Ausführungsform ist vorteilhaft, da hier keine die Rekombinase kodierenden Sequenzen im Genom verbleiben.

Natürlich sind auch Kombinationen der oben beschriebenen Möglichkeiten denkbar.

Bevorzugt ist die Expressionskassette für die Rekombinase auf der Insertionssequenz oder dem Transformationskonstrukt enthalten. Es können auch Masterpflanzen erzeugt werden, die eine Expressionskassette für eine Rekombinase stabil in die plastidäre DNA oder die nukleäre DNA integriert umfassen.

Bevorzugt ist die Rekombinase im Moment des Einbringen des Transformationskonstruktes bzw. der Insertionssequenz in den Plastiden der Masterpflanze funktionell vorhanden.

Besonders bevorzugt wird die Rekombinase gleichzeitig mit oder nach der Einführung der Insertionssequenz in die Plastiden eingebracht bzw. aktiviert. Die Expression bzw. Aktivierung am richtigen Ort und zur richtigen Zeit kann durch verschiedene Ansätze sichergestellt werden:

a) Induzierbare Expression

Die Expression einer Rekombinase kann unter Verwendung eines induzierbaren Promotor, bevorzugt eines chemisch induzierbaren Promotors, gesteuert werden. Dazu kann beispielsweise die für die Rekombinase kodierende Expressionskassette stabil in die plastidäre oder nukleäre DNA einer Masterpflanze transformiert werden. Erfolgt die Transformation in das Kerngenom, so muss - wie oben beschrieben - die subzelluläre Lokalisation durch geeignete PLS-Transitpeptide sichergestellt werden. Kurz vor oder während der Transformation mit der Insertionssequenz oder dem Transformationskonstrukt wird dann in Abhängigkeit von dem gewählten induzierbaren System die Expression der Rekombinase durch Applikation des entsprechenden Induktors eingeschaltet. Dem Fachmann sind verschiedene Verfahren oder Promotoren zur induzierten Expression bekannt. Als Stimulus können chemische Substanzen oder auch physikalische Stimuli wie beispielsweise erhöhte

Temperatur oder Verwundung etc. fungieren. Verschiedene Beispiele sind weiter unten beschrieben.

b) Induzierbare Aktivität

5

Die Rekombinase kann bereits in den Plastiden der Masterpflanze vorliegen, wenn die Aktivität durch geeignete Techniken erst zum gewählten Zeitpunkt etwa durch Zugabe chemischer Substanzen induziert wird. Entsprechende Verfahren sind für sequenzspezifische Rekombinasen beschrieben (Angrand PO et al. (1998) Nucl Acids Res 26(13):3263-3269; Logie C und Stewart AF (1995) Proc Natl Acad Sci USA 92(13):5940-5944; Imai T et al. (2001) Proc Natl Acad Sci USA 98(1):224-228). Bei diesen Verfahren werden Fusionsproteine aus der Rekombinase und der Ligandenbindedomäne eines Steroidhormonrezeptors (z.B. des humanen Androgenrezeptors, oder mutierte Varianten des humanen Estrogenrezeptors wie dort beschrieben) eingesetzt. Die Induktion kann mit Liganden wie beispielsweise Estradiol, Dexamethason, 4-Hydroxytamoxifen oder Raloxifen erfolgen.

20

Die meisten Rekombinasen sind als Multimer (oft als Tetramer) aktiv (Stark et al., 1992, TIG 8: 432-439). Es ist denkbar, die Multimerisierung induzierbar zu gestalten, indem beispielsweise die natürlichen Multimerisierungsdomänen gegen die Bindungsdomäne eines niedermolekularen Liganden ausgetauscht werden. Zugabe eines dimeren Liganden bewirkt dann Dimerisierung des Fusionsproteins. Entsprechende induzierbare Dimerisierungsverfahren als auch die Herstellung der dimeren Liganden sind beschrieben (Amara JF et al. (1997) Proc Natl Acad Sci USA 94(20): 10618-1623; Muthuswamy SK et al. (1999) Mol Cell Biol 19(10): 6845-6857; Schultz LW und Clardy J (1998) Bioorg Med Chem Lett 8(1):1-6; Keenan T et al. (1998) Bioorg Med Chem. 6(8):1309-1335).

35

c) Cotransfektion

Bevorzugt wird die Expressionskassette kodierend für die Rekombinase gleichzeitig mit der Insertionssequenz in die Plastiden eingebracht. Dabei können die Expressionskassette für die Rekombinase und die Insertionssequenz auf einem DNA-Molekül oder auf zwei getrennten vorliegen. Bevorzugt liegen die beiden Sequenzen auf einem DNA-Molekül zusammen vor, so dass die Expressionskassette in dem die Insertionssequenz umfassenden Transformationskonstrukt enthalten ist.

45

25

Dabei wird in einer besonders bevorzugten Ausführungsform, nachdem homoplastome Pflanzen regeneriert wurden, die Sequenz kodierend für die Rekombinase aus dem Plastom wieder entfernt. Dem Fachmann sind dazu verschiedene Verfahren bekannt, die weiter unten im Detail beschrieben sind.

Ein weiterer Gegenstand der Erfindung betrifft Fusionsproteine von Rekombinasen der Familie der Resolvasen/Invertasen mit Plastidlokalisationssequenzen, sowie die für diese kodierenden Nukleinsäuresequenzen. Umfasst sind ferner Expressionskassetten, die besagte Nukleinsäuresequenzen unter Kontrolle eines im pflanzlichen Zellkern funktionellen Promotors enthalten. Entsprechende Promotoren sind dem Fachmann bekannt und weiter unten beschrieben. Bevorzugte Resolvasen/Invertasen sind die Φ C31 Rekombinase, die R4 Rekombinase, das Hin Inversionssystem aus *Salmonella typhimurium* und die TP901-1 Rekombinase, sowie die Rekombinasen Gin des Mu Phagen, Cin des Bakteriophagen P und Pin des ϕ 14 Phagen. Ganz besonders bevorzugt sind die Φ C31 Rekombinase und die TP901-1 Rekombinase. Am meisten bevorzugt sind die Proteine gemäß SEQ ID NO: 41, 42, 46, 47 und 61. Als Expressionskassetten bevorzugt sind insbesondere die Sequenzen gemäß SEQ ID NO: 40 und 45. Die angegebene Sequenz der TP901-1 Rekombinase gemäß SEQ ID NO: 46 und 47 enthält im Vergleich zu der in der GenBank hinterlegten (SEQ ID NO: 61) einen Aminosäureaustausch.

Ein weiterer Gegenstand der Erfindung betrifft Expressionskassetten, die Nukleinsäuresequenzen kodierend für Rekombinasen der Familie der Resolvasen/Invertasen unter Kontrolle eines in pflanzlichen Plastiden funktionellen Promotors enthalten. Entsprechende Promotoren sind dem Fachmann bekannt und weiter unten beschrieben. Bevorzugte Resolvasen/Invertasen sind die Φ C31 Rekombinase, die R4 Rekombinase, das Hin Inversionssystem aus *Salmonella typhimurium* und die TP901-1 Rekombinase, sowie die Rekombinasen Gin des Mu Phagen, Cin des Bakteriophagen P und Pin des ϕ 14 Phagen. Ganz besonders bevorzugt sind die Φ C31 Rekombinase und die TP901-1 Rekombinase.

"Rekombinase-Erkennungssequenz" (infolge "RE-Sequenz") meint allgemein solche Sequenzen, die unter den Bedingungen in den Plastiden der jeweils verwendeten pflanzlichen Zelle oder Pflanze die Erkennung und nachfolgende Verwendung als Rekombinationssubstrat durch eine Rekombinase erlauben.

Bevorzugt ist die Geschwindigkeit einer Rekombinationsreaktion einer bestimmten Rekombinase unter Einsatz einer RE-Sequenz (beispielsweise ihrer natürlichen Erkennungssequenz) mindestens zehn-

26

fach so hoch wie für den Einsatz einer beliebigen anderen nicht-homologen Sequenz, bevorzugt mindestens hundertfach so hoch, besonders bevorzugt mindestens eintausendfach so hoch, am meistens bevorzugt mindestens zehntausendfach so hoch. Dabei meint eine

5 "nicht-homologe Sequenz" bevorzugt solche, die bei gleicher Länge, eine Sequenzidentität zu der bevorzugten Erkennungssequenz (beispielsweise ihrer natürlichen Erkennungssequenz) der jeweiligen Rekombinase von weniger als 90 %, bevorzugt weniger als 50 %, besonders bevorzugt weniger als 30 % aufweisen.

10

Besonders bevorzugt sind RE-Sequenzen der oben beschriebenen sequenzspezifischen Rekombinasen. Bevorzugt ist die RE-Sequenz einer bestimmten Rekombinase singulär in der plastidären DNA, d.h. eine Rekombination wird nur an der so vordefinierten Stelle

15 erzeugt. Es sind jedoch auch Fälle denkbar, bei denen mehr als eine RE-Sequenz im Plastom vorhanden ist. Dies ist insbesondere der Fall, wenn die RE-Sequenz in duplizierten Genen (z.B. in invertierten "Repeats") lokalisiert ist. Im letzteren Fall liegen mehr als eine identische RE-Sequenz vor, ihr Kontext ist jedoch

20 identisch, so dass auch hier eine gezielte Insertion erfolgt. Es ist sogar bevorzugt, dass die Integration in alle Kopien erfolgt, so dass auch eine Insertion in allen Kopien erforderlich ist. RE-Sequenzen, die zwar mehr als einmal in einem Plastom auftreten, jedoch im gleichen plastomen Kontext lokalisiert sind (z.B. in

25 "Repeats" oder Genduplikationen) sind im Rahmen dieser Erfindung unter dem Begriff der singulären RE-Sequenzen subsumiert. Natürlich können singuläre RE-Sequenzen unterschiedlicher Rekombinasen parallel in einem Plastom vorliegen.

30 Anstelle der natürlichen RE-Sequenzen der genutzten Rekombinasen können auch modifizierte Erkennungsregionen genutzt werden, wie sie zum Beispiel für die FLP oder Cre Rekombinase beschrieben wurden (Senecoff JF und Cox MM (1986) J Biol Chem 261:7380-7386; Albert H et al. (1995) Plant J 7:649-659). Bevorzugt werden

35 jedoch die natürlichen Sequenzen im Rahmen dieser Erfindung genutzt. Darüber hinaus können die Rekombinasen selbst durch eine Mutagenese verbessert werden, wie es etwa für die FLP Rekombinase gezeigt wurde (Buchholz F et al. (1998) Nature Biotech 16:657-662).

40

Es ist auch denkbar, zwei nicht miteinander rekombinierende RE-Sequenzen (die von ein und der selben oder verschiedenen Rekombinasen erkannt werden) flankierend zu einem zu integrierenden DNA-Segment zu nutzen. Diese können bevorzugt mit jeweils kompatiblen

45 Erkennungsregionen in einem anderen Molekül rekombiniert werden (Hoess RH et al. (1986) Nucl Acids Res 14: 2287-2300).

27

- Grundsätzlich kann zwischen "wildtyp" RE-Sequenzen und "pseudo" RE-Sequenzen unterschieden werden. Dabei meint "wildtyp" RE-Sequenz solche Sequenzen, wie sie im natürlichen System von der Rekombinase genutzt werden. Beispielsweise die loxP Sequenz für
- 5 die cre Rekombinase oder die FRT Sequenz für die FLP Rekombinase oder die GIX Sequenz für die Gin Rekombinase. Diese Sequenzen können aus ihren homologen Ursprungssystemen (beispielsweise den natürlichen Phagen) abgeleitet und vorteilhaft im Rahmen dieser Erfindung eingesetzt werden. "Pseudo" RE-Sequenzen meint solche
- 10 Sequenzen, die von einer Rekombinase erkannt und als Substrat für eine Rekombination genutzt werden können jedoch Sequenzen haben, die mit der "wildtyp" RE-Sequenz nicht identisch sind. Solche Sequenzen können beispielsweise aus heterologen Organismen isoliert werden, oder mittels künstlicher Mutagenese erzeugt werden.
- 15 Beispielsweise können manche Rekombinasen - wie beispielsweise Φ C31 - in den Genomen mancher Eukaryoten Rekombinationen bewirken, haben also auch hier funktionelle RE-Sequenzen, die ähnlich aber nicht identisch zu der "wildtyp" RE-Sequenz sind. "Pseudo" RE-Sequenzen können mittels Sequenzalignment, Sekundär-
- 20 strukturvergleich, Deletions- oder Punktmutationsanalyse identifiziert werden. Dabei kommen vorteilhafterweise funktionelle Tests zum Einsatz, um die Wirkung der Rekombinase auf die RE-Sequenzen zu bestimmen. Ferner umfasst sind hybride RE-Sequenzen, die sich aus Teilen einzelner RE-Sequenzen - beispielsweise -
- 25 aus jeweils einem Anteil einer "wildtyp" RE-Sequenz und einer "pseudo" Re-Sequenz zusammensetzen.

Die Rekombinase wird in die Plastiden vor, gleichzeitig und/oder nach Einführung des Transformationskonstruktes eingebracht.

30

- Bevorzugt ist die eingesetzte Pflanze oder von dieser abgeleitete Zelle bezüglich der RE-Sequenz überwiegend homoplastom oder homotransplastom, d.h. dass die überwiegende Anzahl der in einem Plastid enthaltenen plastidären DNA-Moleküle diese RE-Sequenz
- 35 aufweisen. Solche Pflanzen werden auch als Masterpflanzen im Rahmen dieser Erfindung bezeichnet.

Prinzipiell können zwei Arten von RE-Sequenzen genutzt werden:

- 40 a) Natürliche, endogene RE-Sequenz

- Eine natürlicherweise im Plastom vorkommende Sequenz kann als Erkennungssequenz für beispielsweise eine chimäre Rekombinase fungieren. Chimäre Rekombinasen können beispielsweise durch
- 45 Fusion einer die Sequenz besagter Erkennungsregion bindenden Zinkfingerdomäne und einer katalytischen Domäne einer Rekombinase bestehen (WO 96/06166). Wie solche Zinkfingerdomänen

bzw. solche chimären Rekombinasen erstellt werden, ist dem Fachmann bekannt (Beerli RR et al. (2000) Proc Natl Acad Sci USA 97(4):1495-1500; Beerli RR et al. (2000) J Biol Chem 275(42):32617-32627; Segal DJ and Barbas CF 3rd. (2000) Curr Opin Chem Biol 4(1):34-39; Kang JS and Kim JS (2000) J Biol Chem 275(12):8742-8748; Beerli RR et al. (1998) Proc Natl Acad Sci USA 95(25):14628-14633; Kim JS et al. (1997) Proc Natl Acad Sci USA 94(8):3616-3620; Klug A (1999) J Mol Biol 293(2):215-218; Tsai SY et al. (1998) Adv Drug Deliv Rev 30(1-3):23-31; Mapp AK et al. (2000) Proc Natl Acad Sci USA 97(8):3930-3935; Sharrocks AD et al. (1997) Int J Biochem Cell Biol 29(12):1371-1387; Zhang L et al. (2000) J Biol Chem 275(43):33850-33860). Verfahren zur Herstellung und Selektion von Zink-Finger DNA-Bindedomänen mit hoher Sequenzspezifität sind beschrieben (WO 96/06166, WO 98/53059, WO 98/53057).

Darüber hinaus ist es denkbar, bekannte Rekombinasen einer Mutagenese zu unterziehen, bis sie eine ausgewählte Sequenz des Plastidengenoms einer betrachteten Pflanzenspezies als Substrat erkennen. Pflanzen mit endogenen, natürlichen RE-Sequenzen stellen quasi natürlich vorkommende "Masterpflanzen" dar. Bei ihnen ist die RE-Sequenz natürlicherweise homoplastom vorhanden. Dies erübrigt die Einführung und Selektion künstlicher RE-Sequenzen.

25
b) Künstlich eingeführte RE-Sequenz

Dem Fachmann ist bewusst, dass die in eine Masterpflanze eingebrachte RE-Sequenz nicht natürlich sein muss. Prinzipiell kann jede RE-Sequenz - beispielsweise eine natürliche, mutierte oder pseudo RE-Sequenz - einer beliebigen Rekombinase an jede beliebige Stelle der plastidären DNA insertiert werden. Die Herstellung erfolgt bevorzugt unter Verwendung eines Konstrukts zur Insertion der RE-Sequenz (infolge RE-Konstrukt).

Bevorzugt umfasst das RE-Konstrukt einen Selektionsmarker, um die zur Erzeugung entsprechender Masterpflanzen erforderliche Selektion transplastomer Pflanzen mit der erfolgreich insertierten RE-Sequenz zu erleichtern. Dem Fachmann sind verschiedene Selektionsmarker bekannt, die eine Selektion von Plastiden ermöglichen (s.u.). Bevorzugt sind aadA, nptII oder BADH, wobei aadA besonders bevorzugt ist. Die Selektion erfolgt beispielsweise mit Hilfe des dem Fachmann bekannten "segregation and sorting" Prozess (beispielhaft unter Beispiel 4 beschrieben). Der Selektionsmarker ist bevorzugt so konstruiert, dass eine nachfolgende Deletion aus dem Plastom

29

ermöglicht wird. Entsprechende Verfahren sind dem Fachmann bekannt und unten beschrieben.

5 Das RE-Konstrukt kann neben dem Selektionsmarker weitere Sequenzen enthalten. Diese können beispielsweise weitere regulatorische Elemente für die Expression der infolge ein-
zuführenden Insertionssequenzen enthalten. Der im Rahmen
des Konstruktes zur Insertion der RE-Sequenz eingeführte
10 Selektionsmarker wird in einer bevorzugten Ausführungsform nach Erhalt der homoplastomen Masterpflanze durch dem Fachmann bekannte Verfahren deletiert (s.u.).

Alternativ, können auch natürliche, endogen am Insertionsort
vorhandene regulatorische Elemente für die Expression in
15 Folge einzubringender Gene genutzt werden.

In einer bevorzugten Ausführungsform umfasst das RE-Konstrukt zur Ermöglichung einer ortsspezifischen Insertion an mindestens einer, bevorzugt an beiden Seiten der RE-Sequenz
20 weitere flankierende Sequenzen (A und B), die eine ausreichende Länge und Homologie zu entsprechenden Zielsequenzen im Plastom (A' und B') haben, um eine ortsspezifische Insertion mittels homologer Rekombination zu gewährleisten.

25 Nicht natürlicherweise in der plastidären DNA vorkommende RE-Sequenzen können auf verschiedene Arten in die plastidäre DNA eingeführt werden. Beispielhaft seien zu nennen:

a) Integration mittels homologer Rekombination (z.B. Doppel-
30 Crossover)

Bevorzugt wird die Integration in das Plastidengenom mit Hilfe der oben beschriebenen, dem Fachmann allgemein bekannten Verfahren der homologen Rekombination (z.B. Doppel-Cross-
35 over) durchgeführt.

b) Integration unter Verwendung natürlicher oder endogener Erkennungssequenzen für ein Enzym geeignet zur Induktion von DNA-Doppelstrangbrüchen (infolge DSB-Erkennungssequenz; siehe
40 unten).

c) Integration in Folge einer Rekombination mit Rekombinasen. Das erfindungsgemäße Verfahren selber kann genutzt werden, um beispielsweise weitere RE-Sequenzen zu integrieren.
45

30

Es sind verschiedene Orte der Lokalisation bzw. Integration einer RE-Sequenz (bei bereits vorhandenen endogenen RE-Sequenzen) bzw. Integration (bei künstlich generierten RE-Sequenzen) für die RE-Sequenz möglich. Beispielhaft seien zu nennen:

5

- a) Lokalisation (Integration) in einer transkriptionell stillen Region

10

Lokalisation (Integration) der RE-Sequenz in einer transkriptionell stillen Region des Plastidengenoms (intergenische Region) ist die bevorzugte Ausführungsform. Eine Störung der plastidären Funktionen kann so weitgehend ausgeschlossen werden. Hierbei ist zu beachten, dass für eine Expression (z.B. von Selektionsmarkern oder anderen Genen von Interesse) gegebenenfalls entsprechende regulatorische Elemente wie Promotoren etc. mit eingebracht werden müssen.

15

- b) Lokalisation (Integration) in eine transkriptionell aktive, aber nicht-kodierende (intercistronische) Region

20

Diese Lokalisation (Integration) hat den Vorteil, dass dadurch die einzubringende Insertionssequenz letztendlich in einem plastidären Operon kodiert ist und Promotor(en) bzw. Terminator(en) nicht gesondert mit eingebracht werden müssen, sondern die endogen an diesem Locus vorhandenen ausgenutzt werden können (aber nicht müssen). Es sollten in einem solchen Fall lediglich Ribosomenbindestellen in geeignetem Abstand stromaufwärts der kodierenden Region der einzubringenden Fremd-Gene vorhanden sein.

30

Es ist jedoch auch denkbar, dass eine intergenische Region nicht vollständig transkriptionell still ist, weil beispielsweise eine nur ineffiziente Termination der Transkription von einem benachbarten Gen oder Operon erfolgt.

35

- c) Lokalisation (Integration) in eine transkriptionell aktive, kodierende Regionen.

40

Die unter a) und b) beschriebene Lokalisation (Integration) der RE-Sequenz an einem nicht kodierenden Locus, hat den Vorteil, dass die Insertion der Fremd-DNA die Funktion des Plastidengenoms mit hoher Wahrscheinlichkeit nicht beeinflusst. Nicht-kodierende Bereiche sind jedoch weniger konserviert als kodierende. Um ein möglichst universelles Verfahren zu haben, dass in vielen Pflanzenarten funktioniert, ist in einer besonders bevorzugten Ausführungsform, die RE-Sequenz (und infolge die Insertionssequenz) in der kodierende Sequenz

45

31

- eines bestehenden Gens lokalisiert. Die Zerstörung der Genfunktion durch die Einführung der RE-Sequenz (bei einer künstlich generierten RE-Sequenz) oder durch die nachfolgende Einführung der Insertionssequenz wird auf erfinderische Weise
- 5 dadurch verhindert, dass die RE-Sequenz bzw. die Insertionssequenz in einer bevorzugten Variante dieser Ausführungsform im Rahmen eines Introns eingebracht wird. Auf diese Weise wird die vollständige kodierende mRNA des Gens, in das die RE-Sequenz inseriert wurde, durch Spleißen der pre-RNA
- 10 wieder generiert. Alternativ kann die Insertion auch in ein nicht-essentielles plastidäres Gen erfolgen. Auch kann das durch die Insertion der RE-Sequenz zerstörte Gen in einer nachfolgenden Transformation wieder eingebracht werden.
- 15 Das RE-Konstrukt hat bevorzugt - insbesondere wenn die Insertion in eine transkriptionell aktive oder gar translatierte Plastomregion erfolgt - die Struktur und Sequenz eines Introns. In der Regel werden dazu natürlicherweise vorkommende Introns so modifiziert, dass sie den Erfordernissen des erfindungsgemäßen Verfahrens genügen. Bevorzugt erfolgt die Insertion derart, dass die
- 20 insertierte Sequenz durch Spleißen der pre-mRNA restlos entfernt wird. Die herausgespleißte RNA (also das artifizielle Intron) stellt dann die mRNA beispielsweise für die Translation auf ihr kodierter Proteine dar. Da die Transkription des Introns der
- 25 regulatorischen Kontrolle des Gens, in das das Intron integriert wurde, unterliegt, kann man sich alle regulatorischen Elemente stromaufwärts bzw. stromabwärts der Gene oder des Gens von Interesse in dem Intron sparen. Dadurch kann man die Konstrukte entsprechend klein halten. Darüber hinaus sind alle Introns nutz-
- 30 bar, wenn man gleichzeitig die entsprechenden Faktoren, die das Spleißen vermitteln, in den Plastiden exprimiert oder sie in diese importiert. Bevorzugt sind die das Spleißen unterstützenden Faktoren im Intron selbst kodiert. Besonders bevorzugt werden in dieser Ausführungsform Introns der Gruppe II, die selbst
- 35 wenigstens einen der für das Spleißen notwendigen Faktoren kodieren. Dazu gehört das Ll.1trB Intron aus *Lactococcus*. Ebenfalls bevorzugt sind Introns, die natürlicherweise in Plastiden höherer Pflanzen vorkommen, besonders Introns der Gruppe II, ganz besonders bevorzugt Introns, die für ein Protein kodieren, am
- 40 meisten bevorzugt Introns der trnK Gene des Plastidengenoms. Im letzteren Fall sind besonders bevorzugt die Introns aus den trnK Genen der Plastiden aus den Arten *Arabidopsis*, Mais und Tabak.
- Bevorzugt sind Introns mit einer selbst-spleißenden Aktivität,
- 45 die nicht von weiteren Proteinfaktoren abhängt oder die allgemeine Faktoren benutzen, die universell und damit auch in

32

Plastiden vorhanden sind. Zu diesen Introns gehören beispielsweise

- a) das Intron der Gruppe I aus *Tetrahymena* (GenBank Acc.-No.: X54512; Kruger K et al. (1982) Cell 31:147-157; Roman J und Woodson SA (1998) Proc Natl Acad Sci USA 95:2134-2139)
- b) das rIl-Intron der Gruppe II aus *Scenedesmus obliquus* (GenBank Acc.-No.: X17375.2 Nukleotide 28831 bis 29438; Holländer V und Kück U (1999) Nucl Acids Res 27: 2339-2344; Herdenberger F et al. (1994) Nucl Acids Res 22: 2869-2875; Kück U et al. (1990) Nucl Acids Res 18:2691-2697).
- c) das Ll.LtrB Intron (GenBank Acc.-No.: U50902 Nukleotide 2854 bis 5345)
- d) das trnK-Intron aus *Arabidopsis* (GenBank Acc.-No.: AP000423 Nukleotide komplementär 1752 bis 4310), Mais (GenBank Acc.-No.: X86563 Nukleotide komplementär 1421 bis 3909) oder Tabak (GenBank Acc.-No.: Z00044 Nukleotide komplementär 1752 bis 4310)

Sowohl artfremde als auch natürlicherweise in den Plastiden der jeweiligen Pflanze vorkommende Introns können genutzt werden. Zur Vermeidung von durch Sequenzduplikation bedingte Instabilitäten sind artfremde Introns - beispielsweise artfremde trnK-Introns - bevorzugt. Natürlicherweise in den Plastiden der jeweiligen Pflanze vorkommende Introns werden in einer bevorzugten Ausführungsform so modifiziert, dass sie zwar ihre Funktion noch erfüllen können, die Sequenzhomologie jedoch geringer als 95 %, bevorzugt 80 %, besonders bevorzugt 70 % zu der Sequenz des Ausgangsintrons ist.

Besonders bevorzugt sind Introns, die natürlicherweise ein DSB-Enzym (insbesondere eine Homing Endonuklease kodieren). Besonders bevorzugt ist das Intron Cp.LSU2 aus *Chlamydomonas pallidostigmatica*, welches das Enzym I-CpaI kodiert (Turmel M et al. (1995) Mol Biol Evol 12:533-545). Bevorzugt sind ferner die Introns der Gruppe II aus den Mitochondrien der Hefe.

In einer bevorzugten Ausführungsform wird die Intronsequenz dem Insertionsort angepasst, so dass sie an diesem Locus spleißen können. Die Anpassung kann bei Gruppe I Introns die *internal guide sequence* (IGS) bzw. bei den Introns der Gruppe II die *exon binding sequence* (EBS) I oder/und II betreffen.

33

Im Fall des trnK Introns aus Mais konnte gezeigt werden, dass eine Editierung (His420Tyr) der entsprechenden mRNA in Plastiden der Gerste erfolgt (Vogel J et al. (1997) J Mol Biol 270:179-187). In einer bevorzugten Ausführungsform wird daher das matK

5 Gen im trnK-Intron aus Mais bereits durch einen entsprechenden His/Tyr-Austausch auf DNA-Ebene modifiziert, so dass eine RNA-Editierung nicht mehr erforderlich ist.

Die Spleißstelle wird im Fall von Gruppe I Introns bestimmt durch

10 die Paarung der IGS mit dem 5' und/oder 3' zum Intron gelegenen Exon des entsprechenden Transkriptes (Lambowitz AM & Belfort M (1993) Annu Rev Biochem 62:587-622). Durch dem Fachmann bekannte Techniken wie PCR oder synthetisches Erstellen von Nukleotid-

15 sequenzen können die IGS beliebiger Gruppe I Introns entsprechend so angepasst werden, dass ein Spleißen an der vordefinierten Insertionsstelle innerhalb der DSB-Erkennungsregion erfolgt. Bevorzugt wird das Intron CpLSU2 aus *C. pallidostigmatica* genutzt, welches für die Homing Endonuklease I-CpaI codiert. Bevorzugt wird ferner das Gruppe I Intron aus *Tetrahymena thermo-*

20 *phila* (GenBank Acc.-No.: V01416 J01235; Nukleotide 53 bis 465). Die natürlicherweise zu findende IGS kann durch den Fachmann bekannte Techniken an die neue Insertionsstelle angepasst werden.

Selbst-spleißende Introns der Gruppe II besitzen eine konser-

25 vierte Struktur und bestehen im allgemeinen aus 6 verschiedenen Domänen. Domäne I beinhaltet die Exon-Bindestellen (EBS1 und EBS2), die beim Spleißvorgang eine Interaktion mit dem 5' vom Intron gelegenen Exon eingehen. Darüber hinaus findet eine Interaktion zwischen der "δ Region" und der "δ' Region" am 3'

30 Exon statt (Lambowitz AM & Belfort M (1993) Annu Rev Biochem 62:587-622; Michel F & Ferat JL (1995) Annu Rev Biochem 64:435-461). Diese Sequenzen können durch dem Fachmann bekannte Techniken wie synthetisches Erstellen der Introns oder geeignete PCR Methoden jeweils derart angepasst werden, dass eine korrekte

35 Wahl der Spleißstellen an dem in der DSB-Erkennungsregion gewählten Insertionsort gewährleistet ist.

Selbstverständlich können über ein RE-Konstrukte auch gleich mehrere RE-Sequenzen für bevorzugt unterschiedliche Rekombinasen

40 eingebaut werden. Das RE-Konstrukt wird dabei bevorzugt in der Art und Weise erstellt, dass nach der Integration in das Plastom keine Sequenzen dupliziert vorliegen (keine Homologien zu nativen Sequenzbereichen).

45 Bevorzugt wird so zunächst eine bezüglich der insertierten RE-Sequenz überwiegend homotransplastome Pflanze erzeugt, die eine RE-Sequenz in allen oder der überwiegenden Anzahl der Plastiden

der betrachteten Pflanze besitzen. Solche Pflanzen können vorteilhaft als Masterpflanzen eingesetzt werden. Auch wenn der Aufwand zur Insertion einer künstlichen RE-Sequenz in die plastidäre DNA relativ hoch ist und im Fall der Insertion mittels homologer

- 5 Rekombination (z.B. Crossover) dem der zur Zeit im Stand der Technik beschriebenen Verfahren zur Plastidentransformation entspricht, so muss dieser Aufwand lediglich einmalig betrieben werden. Die erhaltene homotransplastome Masterpflanze kann dann für beliebig viele unterschiedliche nachfolgende Transformationen
- 10 unter Verwendung des erfindungsgemäßen Verfahrens eingesetzt werden, was eine erhebliche Steigerung der Transformations-effizienz ermöglicht: Statt den herkömmlichen Selektionsprozess für eine homotransplastome Pflanze jedes Mal vollständig durchlaufen zu müssen, muss er hier nur einmal realisiert werden.

15

Aufgrund der Vielzahl der im Stand der Technik beschriebenen Rekombinasen mit definierten RE-Sequenzen, ist es möglich und bevorzugt, Masterpflanzen zu erzeugen, die mehrere verschiedene singuläre RE-Sequenzen in ihr plastidäres Genom eingebaut haben.

- 20 Es können dann auch verschiedene Transformationsvektoren - auch gleichzeitig - eingebracht werden. Beim Nutzen verschiedener Erkennungsregionen, die entweder schon in der Masterpflanze vorhanden sind oder aber jeweils mit dem neu transformierten Transformationskonstrukt eingebracht werden, kann das Verfahren
- 25 beliebig oft wiederholt werden. Nicht alle eingebrachten Vektoren müssen mittels Rekombinasen insertiert werden (es können auch verschiedene Insertionssequenzen mittels verschiedener Integrationsmechanismen insertiert werden); nicht alle - optional kein - Transformationskonstrukte müssen einen Selektionsmarker
- 30 kodieren. Selektionsmarker können auch in trans auf einem anderen Konstrukt vorhanden sein. Dieses zusätzliche Konstrukt kann auch nur einen Selektionsmarker umfassen (beispielsweise "binding type marker" wie z.B. eine Punktmutation in der 16SrDNA zum Erreichen einer Spectomycinresistenz).

35

Für im Rahmen der Erfindung vorteilhaft einzusetzende RE-Sequenzen seien beispielhaft - aber nicht einschränkend - die in nachfolgender Tabelle 1 wiedergegebenen Erkennungssequenzen für die jeweiligen aufgeführten Rekombinasen genannt.

40

45

Tabelle 1: Erkennungssequenzen von Rekombinasen

	Rekombinase	Referenz	RE-Sequenz
5	Φ C31 attP	Groth AC et al. (2000) Proc Natl Acad Sci USA 97: 5995-6000	GTAGTGCCCCAACTGGGGTAAC- CTTTGAGTTCTCTCAGTTGGGGGCGTAG
10	Φ C31 attB	Groth AC et al. (2000) Proc Natl Acad Sci USA 97: 5995-6000	CGGGTGCCAG- GGCGTGCCCTTGGGCTCCCCGGGCGCGTAC
	TP901-1 attP	Brondsted L (1999) Appl Env Microbiol 65: 752-758	TCCAACCTCGCTTAATTGCGAGTTTTTATTTTCG TTTATTTCAATTAAGGTAATAAAAACTCCT TTTAAGGAGTTTCTG
15	TP901-1 attB	aus Y15043 Nu- keotide (kompli- mentär 1925-1825)	AGCTTTGGCAAAAAAGCAAAAAGCATTTCAC TTGATTGAGATGTAAATTGTGTGGCAATTAT CAGTATTTTAATTTTGCTTTTGTGCCAAATT TGATA
20	Cre		ATAACTTCGTATA GCATACAT TATAC- GAAGTTAT (loxP nativ)
	Cre	Albert H et al, 1995, Plant J: 7: 649-659	lox66: ataacttcgtata gcatacat tatacgaacggta
	Cre	"	lox71: taccgttcgtata gcatacat tatacgaagttat
25	Cre	"	lox76: ataacttcgtata gcatacat tatacgcccgga
	Cre	"	lox75: taccgggctgata gcatacat tatacgaagttat
30	Cre	"	lox43: ataacttcgtata gcatacat tataggtaccgag
	Cre	"	lox44: aatgcatgctata gcatacat tatacgaagttat
	FLP		GAAGTTCCTATaC TTTCTAGA GAATAG- GAACTTC C GAATAGGAACTTC
35	FLP		GAAGTTCCTATaC TTTCTAGA GAATAG- GAACTTC
40	λ Integrase attP		aatgctctgttacaggtcactaataccatcta agtagttgattcatagtgactgcatatgttgt gttttacagtattatgtagtctgtttttatg caaatctaatttaatatattgatatttatat cattttacgtttctcgttcagcttttttatac taagttggcattataaaaaagcattgcttatc aatttggtgcaacgaacaggtcactatcagtc aaaataaaatcattatttgatttc
45	λ Integrase attB		aagcctgcttttttataactaacttgag

Rekombinase	Referenz	RE-Sequenz
R4 attB	Olivares EC (2001) Gene 278: 167-176	agttgcccattgaccatgccgaagcagtggtga- gaagggcaccggcagacac
5 R4 attP	Olivares EC (2001) Gene 278: 167-176	gcatgttccccaagcgataccacttgaagca gtgggtactgcttgtgggtacactctgcgggtg

Dabei sind auch Abweichungen (Degenerationen) der Erkennungs-
 10 sequenz umfasst, die dennoch Erkennung und Rekombination durch
 die jeweilige Rekombinase ermöglichen, beispielsweise pseudo RE-
 Sequenzen. Ferner sind Kernsequenzen ("Core"-Sequenzen) dieser
 Erkennungssequenzen umfasst. Es ist bekannt, dass auch die
 inneren Anteile der Erkennungssequenzen für eine Rekombination
 15 genügen und dass die äußeren nicht unbedingt relevant sind,
 jedoch die Effizienz der Rekombination mitbestimmen können.
 Ferner können die angegebenen Sequenzen beliebig an beiden
 Seiten ergänzt oder erweitert werden, indem beispielsweise,
 größere flankierende Bereiche aus den entsprechenden natürlichen
 20 Sequenzen übernommen werden. Der Begriff der RE-Sequenz umfasst
 insofern auch alle wesentlichen gleichen Erkennungssequenzen. Im
 wesentlichen gleiche Erkennungssequenzen meint solche Erkennungs-
 sequenzen, die zwar Abweichungen von der für das jeweilige Enzym
 als optimal gefundenen Erkennungssequenz aufweisen, jedoch eine
 25 Rekombination durch dasselbe noch erlauben.

Da die loxP Sequenz bislang nur im P1 Phagengenom gefunden wurde,
 bedingt die Verwendung der identischen loxP Sequenz eine vorher-
 gehende Einführung derselben in das Plasmid (s.u. mittels z.B.
 30 homologer Rekombination). Verschiedene Studien haben jedoch
 bereits gezeigt, dass eine exakte Identität mit der natürlichen
 loxP Sequenz nicht zwingend für eine Cre-vermittelte Rekombi-
 nation erforderlich ist (Sternberg et al. (1981) J Mol Biol
 150:487-507; Sauer J (1992) Mol Biol 223:911-928; Sauer (1996)
 35 Nucl Acids Res 24:4608-4613).

Orientierung und Lokalisation der RE-Sequenzen

Deletion: Für eine Deletion werden zwei RE-Sequenzen, die durch
 40 die Einwirkung einer Rekombinase miteinander rekombinieren
 können, intramolekular mit gleicher Orientierung plaziert.
 Die zu deletierende Sequenz befindet sich zwischen den beiden
 RE-Sequenzen. Es entstehen 2 Moleküle und jedes trägt eine
 der beiden Sequenzen, die sich aus der Rekombination besagter
 45 RE-Sequenzen ergeben.

37

Insertion: Die Insertion ist die Rückreaktion der Deletion. Zwei miteinander rekombinierbare RE-Sequenzen werden auf 2 Molekülen plaziert. Das zu insertierende Molekül ist in der einfachsten Anwendung zirkulär. Die Integration wird in der Art und Weise
5 vorgenommen, in der sich beide RE-Sequenzen in entsprechender Orientierung anordnen. Das entstehende Produkt trägt dann zwei Sequenzen, die sich aus der Rekombination besagter RE-Sequenzen ergeben, in entsprechender Orientierung.

- 10 Es ist auch möglich, zwei RE-Sequenzen, die nicht miteinander rekombinieren können, auf einem Molekül mit gleicher Orientierung zu plazieren. In dem zweiten Molekül, welches dann nicht zwangsläufig zirkulär sein muss, werden dann ebenfalls zwei RE-Sequenzen, welche nicht miteinander aber mit denen auf dem anderen
15 Molekül rekombinieren können, mit gleicher Orientierung plaziert (s. unten "Replacement" Strategie). Das Ergebnis der bevorzugten Reaktion ist ein Austausch der zwischen den RE-Sequenzen auf beiden Molekülen befindlichen Sequenzen.

- 20 Inversion: Zwei miteinander rekombinierbare RE-Sequenzen werden auf einem Molekül mit gegenläufiger Orientierung plaziert. Die zugehörige Rekombinase kann dann eine Inversion der zwischen den RE-Sequenzen befindlichen Sequenzen bedingen. Sofern die durch die Rekombination entstehenden hybriden RE-Sequenzen nicht
25 durch Einwirkung besagter Rekombinase allein wieder miteinander rekombinieren können, ist die Inversion irreversibel.

- "Enzym geeignet zur Induktion von DNA-Doppelstrangbrüchen an der Erkennungssequenz zur gezielten Induktion von DNA-Doppelstrang-
30 brüchen" (infolge "DSBI-Enzym" für "double strand-break inducing enzyme") meint allgemein all solche Enzyme, die in der Lage sind, sequenzspezifisch Doppelstrangbrüche in doppelsträngige DNA zu erzeugen. Beispielsweise aber nicht einschränkend sind zu nennen:

35

1. Restriktionsendonukleasen, bevorzugt Typ II Restriktionsendonukleasen, besonders bevorzugt Homing-Endonukleasen wie weiter unten im Detail beschrieben.
- 40 2. Künstliche Nukleasen wie weiter unten im Detail beschrieben, wie beispielsweise chimere Nukleasen, mutierte Restriktions- oder Homing-Endonukleasen oder RNA-Proteinpartikel abgeleitet von mobilen Introns der Gruppe II.
- 45 Sowohl natürliche als auch künstlich hergestellte DSBI-Enzyme sind geeignet. Bevorzugt sind all solche DSBI-Enzyme, deren Erkennungssequenz bekannt ist und die entweder in Form ihrer

Proteine (beispielsweise durch Aufreinigung) gewonnen oder unter Verwendung ihrer Nukleinsäuresequenz exprimiert werden können.

- Als künstliche DSBI-Enzyme seien beispielhaft chimäre Nukleasen zu nennen, die sich aus einer unspezifischen Nukleasedomäne und einer sequenzspezifischen DNA-Bindungsdomäne (z.B. bestehend aus Zinkfingern) zusammensetzen (Smith J et al. (2000) Nucl Acids Res 28(17):3361-3369; Bibikova M et al. (2001) Mol Cell Biol. 21:289-297; Chandrasegaran S und Smith J (1999) Biol Chem 380:841-848; Kim YG und Chandrasegaran S (1994) Proc Natl Acad Sci USA 91:883-887; Kim YG et al. (1996) Proc Natl Acad Sci USA 93:1156-1160; Nahon E und Raveh D (1998) Nucl Acids Res 26:1233-1239).
- Bevorzugt wird das DSBI-Enzym unter Kenntnis seiner spezifischen Erkennungssequenz so ausgewählt, dass es zusätzlich zu der Ziel-Erkennungssequenz keine weitere funktionellen Erkennungsregionen im Plastom besitzt. Ganz besonders bevorzugt sind deshalb Homing-Endonukleasen (Übersicht: Belfort M und Roberts RJ (1997) Nucleic Acids Res 25:3379-3388; Jasin M (1996) Trends Genet 12:224-228; Internet: <http://rebase.neb.com/rebase/rebase.homing.html>; Roberts RJ and Macelis D (2001) Nucleic Acids Res 29: 268-269). Diese erfüllen aufgrund ihrer langen Erkennungssequenzen diese Anforderung. Aufgrund der geringen Größe des Plastoms ist es jedoch auch denkbar, dass DSBI-Enzyme mit kürzeren Erkennungssequenzen (beispielsweise Restriktionsendonukleasen) mit Erfolg eingesetzt werden können.

Ganz besonders bevorzugt sind

- 30 - I-NanI (Elde M et al. (1999) Eur J Biochem. 259:281-288; GenBank Acc.-No.: X78280, Nukleotide 418 bis 1155),
- I-NitI (GenBank Acc.-No.: X78277, Nukleotide 426 bis 1163),
- 35 - I-NjaI (GenBank Acc.-No.: X78279, Nukleotide 416 bis 1153),
- I-PpoI kodiert auf der extrachromosomaler DNA im Kern von *Physarum polycephalum* (Muscarella DE und Vogt VM (1989) Cell 56:443-454; Lin J und Vogt VM (1998) Mol Cell Biol 18:5809-5817; GenBank Acc.-No.: M38131, Nukleotide 86 bis 577; siehe auch SEQ ID NO: 4),
- 40 - I-PspI (GenBank Acc.-No.: U00707, Nukleotide 1839 bis 3449),

39

- I-ScaI (Monteilhet C et al. (2000) Nucleic Acids Res 28: 1245-1251; GenBank Acc.-No.: X95974, Nukleotide 55 bis 465)
- Endo SceI (Kawasaki et al. (1991) J Biol Chem 266:5342-5347, 5 identisch zu F-SceI; GenBank Acc.-No.: M63839, Nukleotide 159 bis 1589),
- I-SceII (Sarguiel B et al. (1990) Nucleic Acids Res 18:5659-5665),
- 10 - I-SceIII (Sarguiel B et al. (1991) Mol Gen Genet. 255:340-341),
- I-Ssp6803I (GenBank Acc.-No.: D64003, Nukleotide 35372 bis 15 35824),
- I-TevI (Chu et al. (1990) Proc Natl Acad Sci USA 87:3574-3578; Bell-Pedersen et al. (1990) Nucleic Acids Res 18:3763-3770; GenBank Acc.-No.: AF158101, Nukleotide 20 144431 bis 143694),
- I-TevII (Bell-Pedersen et al. (1990) Nucleic Acids Res 18:3763-3770; GenBank Acc.-No.: AF158101, Nukleotide 45612 bis 44836),
- 25 - I-TevIII (Eddy et al. (1991) Genes Dev. 5:1032-1041),

Ganz besonders bevorzugt sind kommerziell erhältliche Homing-Endonukleasen wie I-PpoI, PI-PspI oder PI-SceI. Am meisten bevor- 30 zugt ist I-PpoI.

Am meisten bevorzugt ist die Homing-Endonuklease gemäß den Proteinsequenzen beschrieben durch SEQ ID NO: 5. Bei der Herstellung entsprechender Expressionskassetten werden demzufolge 35 Nukleinsäuresequenzen eingesetzt, die für ein Protein gemäß SEQ ID NO: 5 kodieren, insbesondere bevorzugt ist dabei die Verwendung der Nukleinsäuresequenzen gemäß SEQ ID NO: 4.

Weiterhin können Restriktionsendonukleasen (Typ II) vorteil- 40 haft als DSBI-Enzyme verwendet werden, wenn sie keine Schnittstellen im Plastom aufweisen. Besonders bevorzugt sind dabei die Restriktionsendonukleasen NotI, CciNI, FseI, SbfI, SdaI und Sse8387I.

45 Besonders bevorzugt werden DSBI-Enzyme entsprechend den oben für Rekombinasen beschriebenen Verfahren so exprimiert, dass die Ausübung ihrer Funktion in den Plastiden gewährleistet ist.

40

Besonders bevorzugt ist dabei die Expression als Fusionsprotein mit einer PLS oder die Integration einer Expressionskassette in die plastidäre DNA.

- 5 "Erkennungssequenz zur gezielten Induktion von DNA-Doppelstrangbrüchen" (infolge "DSB-Erkennungssequenz" für Doppelstrangbruch-Erkennungssequenz) meint allgemein solche Sequenzen, die unter den Bedingungen in den Plastiden der jeweils verwendeten pflanzlichen Zelle oder Pflanze die Erkennung und Spaltung durch ein
- 10 DSBI-Enzym erlauben. Besonders bevorzugt sind DSB-Erkennungssequenzen für Homing-Endonukleasen, die natürlicherweise in Mitochondrien oder Kern anderer Organismen kodiert sind. Es können auch DSB-Erkennungssequenzen der Homing-Endonukleasen genutzt werden, die aus Plastiden (beispielsweise von Grünalgen) stammen.
- 15 Bevorzugt ist die DSB-Erkennungssequenz singulär in der plastidären DNA, d.h. ein Doppelstrangbruch wird nur an der vordefinierten Stelle erzeugt. Es sind jedoch auch Fälle denkbar, bei denen mehr als eine DSB-Erkennungssequenz im Plastom vorhanden ist. Dies ist insbesondere der Fall, wenn die DSB-Erkennungs-
- 20 sequenz in duplizierten Genen (z.B. in invertierten "Repeats") lokalisiert ist. Im letzteren Fall liegen mehr als eine identische DSB-Erkennungssequenz vor, ihr Kontext ist jedoch identisch. Es ist sogar bevorzugt, dass die DSB-Erkennungssequenz in alle Kopien der "Repeats" vorhanden ist, so dass auch ein Schnitt in
- 25 allen Kopien erfolgen kann. DSB-Erkennungssequenzen, die zwar mehr als einmal in einem Plastom auftreten, jedoch im gleichen plastomen Kontext lokalisiert sind (z.B. in "Repeats" oder Genduplikationen) sind im Rahmen dieser Erfindung unter dem Begriff der singulären DSB-Erkennungssequenzen subsumiert.
- 30 Bevorzugt ist die eingesetzte Pflanze oder von dieser abgeleitete Zelle bezüglich der DSB-Erkennungssequenz überwiegend homoplastom oder homotransplastom, d.h. dass die überwiegende Anzahl der in einem Plastid enthaltenen plastidären DNA-Moleküle diese DSB-
- 35 Erkennungssequenz aufweisen. Solche Pflanzen werden auch als Masterpflanzen im Rahmen dieser Erfindung bezeichnet.

Tabelle 2: Erkennungssequenzen und Herkunftsorganismus von DSBI-Enzymen ("^^" gibt innerhalb einer Erkennungssequenz die Schnitt-

40 stelle des DSBI-Enzyms an.)

	DSBI-Enzym	Herkunfts- organismus	Erkennungssequenz
5	I-SceII	Saccharomy- ces cerevi- siae	5'-TTTTGATTCTTTGGTCACCC^TGAAGTATA 3'-AAACTAAGAAACCAG^TGGGACTTCATAT
	I-SceIII	Saccharomy- ces cerevi- siae	5'-ATTGGAGGTTTGGTAAC^TATTTATTACC 3'-TAACCTCCAAAACC^ATTGATAAATAATGG
10	F-SceI	Saccharomy- ces cerevi- siae	5'-GATGCTGTAGGC^ATAGGCTTGTT 3'-CTACGACA^TCCGTATCCGAACCAA
	F-SceII	Saccharomy- ces cerevi- siae	5'-CTTTCGCAACA^GTAAAATT 3'-GAAAGGCG^TTGTCATTTTAA
15	I-NanI	Naegleria andersoni	5'-AAGTC'TGGTGCCA^GCACCCGC 3'-TTCAGACC^ACGGTCGTGGGCG
	I-NitI	Naegleria italica	5'-AAGTC'TGGTGCCA^GCACCCGC 3'-TTCAGACC^ACGGTCGTGGGCG
20	I-PpoI	Physarum polycephalum	5'-TAAC TATGACTCTCTTAA^GGTAGCCAAAT 3'-ATTGATACTGAGAG^AATTCCATCGGTTTA "Core Sequenz": CTCTCTTAA^GGTAGC GAGAG^AATTCCATCG
	I-ScaI	Saccharomy- ces capensis	5'-TGTCACATTGAGGTGCACT^AGTTATTAC 3'-ACAGTGTAACTCCAC^GTGATCAATAATG
25	I-Ssp6803I	Synechocy- stis species	5'-GTCGGGCT^CATAACCCGAA 3'-CAGCCCGAGTA^TTGGGCTT
	PI-PspI	Pyrococcus sp.	5'-AAAATCCTGGCAAACAGCTATTAT^GGGTAT 3'-TTT TAGGACCGTTTGTGAT^AATACCCATA
30	I-TevI	Bacterio- phage T4	5'-AGTGGTATCAAC^GCTCAGTAGATG 3'-TCACCATAGT^TGCGAGTCATCTAC
	I-TevII	Bacterio- phage T4	5'-GCTTATGAGTATGAAGTGAACACGT^TATTC 3'-CGAATACTCATACTTCACTTGTG^CAATAAG
35	NotI	Nocardia otitidis- caviarum	5'-GC^GGCCGC-3'
	CciNI	Curto- bacterium citreum N	5'-GC^GGCCGC-3'
40	FseI	Frankia species Eul1b	5'-GGCCGG^CC-3'
	SbfI	Streptomyces species Bf-61	5'-CCTGCA^GG-3'
45	SdaI	Streptomyces diastaticus Ng7-324	5'-CCTGCA^GG-3'
	Sse8387I	Streptomyces specied 8387	5'-CCTGCA^GG-3'

Aufbau des Transformationskonstruktes mit der Insertionssequenz

Auf dem Transformationskonstrukt, welches zur Transformation der Plastiden besagter Masterpflanzen genutzt wird, befindet sich wenigstens eine RE-Sequenz R2, welche unter Einwirkung der Rekombinase mit einer RE-Sequenz R1 in der Masterpflanze rekombinieren kann.

- In einer bevorzugten Ausführungsform führt die Rekombination von R1 mit R2 zu hybriden Rekombinaseerkennungsssequenzen (R1/2 bzw. R2/1), die selber für die Rekombinase allein nur schlechte oder keine Rekombinationssubstrate darstellen. Dabei können besagte hybride Sequenzen durchaus Substrate für die betreffende Rekombinase darstellen, wenn weitere Faktoren hinzutreten (z.B. die Faktor xis und IHF bei Einsatz der λ Rekombinase). Bevorzugt sind diese weiteren Faktoren jedoch nicht gleichzeitig mit der Rekombinase in dem pflanzlichen Organismus vorhanden. Bevorzugt ist die Geschwindigkeit einer Rekombinationsreaktion unter Einsatz der RE-Sequenzen R1 und/oder R2 mindestens zehnfach so hoch wie für den Einsatz der hybriden Sequenzen R1/2 und/oder R2/1, bevorzugt mindestens hundertfach so hoch, besonders bevorzugt mindestens eintausendfach so hoch, am meistens bevorzugt mindestens zehntausendfach so hoch. Dadurch liegt das Gleichgewicht der Rekombination stark auf der Seite der Insertion, so dass die Reaktion quasi irreversibel ist.

In einer bevorzugten Ausführungsform umfasst die RE-Sequenz (beispielsweise die RE-Sequenz R1, R1', R2 oder R2') - in 5'/3'-Richtung - eine erste Sequenz R-5', gefolgt von einer Kernsequenz sowie einer zweiten Sequenz R-3'. Eine besonders bevorzugte Form dieser Sequenzen umfasst die loxP Sequenz, die FRT Sequenz oder die attB bzw. attP der Φ C31 oder TP901-1 Rekombinase.

- In einer weiteren bevorzugten Ausführungsform ist mindestens eine der beiden RE-Sequenzen eine pseudo RE-Sequenz oder eine hybride RE-Sequenz.

Ist die Rekombinase beispielsweise ausgewählt aus der Gruppe der Resolvasen/Invertasen wie - besonders bevorzugt - Φ C31 Rekombinase, TP901-1 Rekombinase oder R4 Rekombinase, so sind die Sequenzen R1 und R2 bevorzugt nicht identisch. Bevorzugt entsprechen die Sequenzen R1 und R2 den Sequenzen attB und attP, oder umgekehrt. Weiterhin kann mindestens eine der beiden Sequenzen attB oder attP durch eine entsprechende pseudo-attB oder pseudo-attP RE-Sequenz ausgetauscht sein.

43

Die attB Sequenz umfasst bevorzugt - in 5'/3'-Richtung - eine erste DNA-Sequenz (attB-5'), gefolgt von einer Kernregion und einer zweiten Sequenz (attB-3').

- 5 Die attP Sequenz umfasst bevorzugt - in 5'/3'-Richtung - eine erste DNA-Sequenz (attP-5'), gefolgt von einer Kernregion und einer zweiten Sequenz (attP-3').

- Besonders bevorzugt ist eine Ausführungsform, bei der durch die
- 10 Rekombinase-vermittelte Rekombination zwischen attB und attP ein Rekombinationsprodukt entsteht, dass nicht mehr als Substrat für die Rekombinase allein (d.h ohne Mitwirkung zusätzlicher Faktoren) fungieren kann. Dieses Rekombinationsprodukt enthält bevorzugt eine Region bestehend aus - in 5'/3'-Richtung - einer ersten
- 15 DNA-Sequenz (attB-5'), gefolgt von dem Rekombinationsergebnis der beiden Kernregionen und einer zweiten Sequenz (attP-3') sowie eine zweite Region bestehend aus - in 5'/3'-Richtung - einer ersten DNA-Sequenz (attP-5'), gefolgt von dem Rekombinationsergebnis der beiden Kernregionen und einer zweiten Sequenz
- 20 (attB-3').

- Insbesondere bevorzugt, ist mindestens eine der beiden RE-Sequenzen attB und/oder attP eine pseudo-RE-Sequenz (att ϕ) umfassend - in 5'/3'-Richtung - eine erste DNA-Sequenz (att ϕ -5'),
- 25 gefolgt von einer Kernregion und einer zweiten Sequenz (att ϕ -3'). Die bei der Rekombination entstehenden Sequenzen umfassen beispielhaft eine Region bestehend aus - in 5'/3'-Richtung - einer ersten DNA-Sequenz (att ϕ -5'), gefolgt von dem Rekombinationsergebnis der beiden Kernregionen und einer zweiten Sequenz
- 30 (attP-3') sowie eine zweite Region bestehend aus - in 5'/3'-Richtung - einer ersten DNA-Sequenz (attP-5'), gefolgt von dem Rekombinationsergebnis der beiden Kernregionen und einer zweiten Sequenz (att ϕ -3'), wobei die entstandenen Regionen nicht mehr als Substrat für die Rekombinase fungieren können, wodurch die
- 35 Insertion quasi irreversibel ist. Alternativ können natürlich auch attB und att ϕ entsprechend rekombiniert werden.

- So führt beispielsweise bei Verwendung der Cre Rekombinase die Rekombination der RE-Sequenzen lox66 und lox71 (s. oben)
- 40 zu einer wildtyp loxP RE-Sequenz und der lox72 RE-Sequenz (5'-taccgttcgtata gcatacat tatacgaacggta-3'), die nur noch sehr ineffizient miteinander rekombinieren können (Araki K et al. (1997) Nucl Acids Res 25:868-872). Ebenso führt die Rekombination der RE-Sequenzen lox76 und lox75 (s. oben) zu einer wildtyp loxP
- 45 RE-Sequenz und der lox72 RE-Sequenz (5'-taccgggctgata gcatacat tatacggccggta-3'), die nur noch sehr ineffizient miteinander rekombinieren können. Gleichermäßen führt die Rekombination der

- RE-Sequenzen lox43 und lox44 zu Wildtyp loxP RE-Sequenz und der lox65 RE-Sequenz (5-aatgcatgctata gcatacat tataggtaccgag-3'), die nur noch sehr ineffizient miteinander rekombinieren können (Albert H et al. (1995) Plant J 7:649- 659). Allgemein sollten
- 5 für die Verwendung im Rahmen dieser Erfindung "right element (RE) mutant lox sites" und "left element (LE) mutant lox sites" etabliert werden, welche miteinander rekombinieren können. Die resultierende loxP (Wildtyp) RE-Sequenz und die RE/LE mutierte RE-Sequenz rekombinieren nur noch sehr schlecht miteinander.
- 10 Dies liegt vor allem daran, dass die LE/RE Mutante nicht mehr als Substrat erkannt werden kann.

- Als Rekombinasen, die für nahezu irreversible Rekombination genutzt werden könnten, sind insbesondere die TP901-1 Rekombi-
- 15 nase, die Φ C31 Rekombinase, die R4 Rekombinase. Möglich sind ferner λ -Integrase (bevorzugt mit Coexpression von IHF), Cre (mit entsprechend mutierten RE-Sequenzen, s.o.), Flp (mit entsprechend mutierten RE-Sequenzen) und die L5 Rekombinase.

- 20 Ferner kann das Transformationskonstrukt einen Selektionsmarker enthalten. Falls in der Masterpflanze bereits ein Selektionsmarker vorhanden ist, unterscheiden sich vorzugsweise besagten Selektionsmarker. Darüberhinaus umfasst das Transformations-
- konstrukt bevorzugt ein oder mehrere Gene von Interesse.

- 25 Regulatorische Elemente können ebenfalls auf dem Transformationskonstrukt vorhanden sein, mit dem RE-Konstrukt bereits in die Masterpflanze eingebracht worden sein, oder es können endogene, natürlicherweise am Integrationsort vorhandene regulatorische
- 30 Elemente genutzt werden.

Das Transformationskonstrukt ist bevorzugt in einem Transformationsvektor enthalten, der weitere Elemente, wie beispielsweise für eine Replikation in Bakterien, umfassen kann.

- 35 Enthält das Transformationskonstrukt nur eine einzelne RE-Sequenz R2, so stellt das Transformationskonstrukt bevorzugt ein zirkuläres DNA-Molekül dar. In diesem Fall kann das Transformationskonstrukt mit dem Transformationsvektor identisch sein. Es ist
- 40 jedoch nicht notwendig, dass die vektoriellen Sequenzen (z.B. Replikationsursprung zur Vermehrung der DNA in Bakterien und ggf. einen bakteriellen Selektionsmarker ebenfalls) in das Plastidengenom integriert werden. Bevorzugt werden diese Sequenzen des Vektorrückgrats vor der Transformation entfernt. Dazu kann man
- 45 beispielsweise die Insertionssequenz von zwei Erkennungsstellen für Restriktionsenzyme umgeben. Nach Vervielfältigen des entsprechenden Vektors in einem Bakterium wie *E. coli* kann man die

45

- isolierte DNA sodann mit der entsprechenden Restriktionsendonuklease bzw. den entsprechenden Restriktionsendonukleasen *in vitro* behandeln. Damit wird die Insertionssequenz vom Vektorrückgrat getrennt und kann *in vitro* durch den Einsatz einer
- 5 Ligase wieder zirkularisiert werden. Die isolierte und zirkularisierte Insertionssequenz kann dann für die Transformation eingesetzt werden. Alternativ kann die Insertionssequenz auch mittels Polymerasekettenreaktion unter Wahl geeigneter Primer amplifiziert und zirkularisiert werden.
- 10 Alternativ und besonders bevorzugt wird das Vektorrückgrat von der Insertionssequenz mit Hilfe von Rekombinasen entfernt. Dazu umgibt man das zu transformierende Fragment mit zwei Erkennungsstellen für eine Rekombinase, mit deren Hilfe man eine Excision
- 15 des zu transformierenden Fragmentes vornehmen kann (beispielsweise Cre-lox, FLP/frt). Es ist möglich, die Ansätze in der Art zu gestalten, dass durch die Aktion der Rekombinase erst die für die Integration der Insertionssequenz notwendige RE-Sequenz R2 entsteht.
- 20 Es ist auch möglich, den kompletten Transformationsvektor zu transformieren und mittels Rekombinasen zu integrieren. Bevorzugt werden in diesem Fall die Vektorrückgratsequenzen anschließend durch geeignete Excisionsverfahren wieder eliminiert.
- 25 Um die Effizienz des erfindungsgemäßen Verfahrens der Integration mittels Erkennungsstellen-spezifischer Rekombination weiter zu steigern, kann beispielsweise die Kopiezahl der zu transformierenden DNA Moleküle erhöht werden, indem das Transformations-
- 30 konstrukt Elemente (z.B. einen plastidären ORI *origin of replication*, Replikationsursprung) umfasst, die es ihm ermöglichen, vor der Integration in die plastidäre DNA im Plastid autonom zu replizieren oder stabil als extrachromosomales DNA-Molekül in den Plastiden zu existieren. Entsprechende Verfahren sind dem Fach-
- 35 mann bekannt (US 5,693,507; US 5,932,479; WO 99/10513). Dieses Verfahren ist bevorzugt, da es die zur Integration im Plastid zur Verfügung stehende Kopienzahl der Insertionssequenzen steigert.
- Alternativ kann auch eine "Replacement" Strategie genutzt werden.
- 40 Dazu werden in den Masterpflanzen zwei RE-Sequenzen R1 und R1', die unter Einwirkung der entsprechenden Rekombinase bzw. Rekombinasen nicht miteinander rekombinieren können, eingebracht (bzw. Sequenzen genutzt, die natürlicherweise dort vorhanden sind). Das Transformationskonstrukt enthält dann Erkennungssequenzen R2 und
- 45 R2', die nicht miteinander unter Einwirkung der Rekombinase oder Rekombinasen rekombinieren, wobei aber R1 mit R2 und R1' mit R2' rekombinieren kann. Dadurch wird die Sequenz zwischen R1 und R1'

- in der Masterpflanze durch die Sequenz zwischen R2 und R2' auf dem Transformationskonstrukt ausgetauscht. Die "Replacement" Strategie ermöglicht es nicht nur, Vektorrückgrat Insertionen zu vermeiden, ermöglicht es darüber hinaus, die Insertionssequenz
- 5 (zuzgl. der Anteile der Erkennungssequenzen R2 und R2') als linearisiertes Fragment einzubringen. Eine solche "Replacement" Strategie wurde zum Beispiel in Mauszellen mit Hilfe unterschiedlicher Erkennungsregionen für die Cre Rekombinase demonstriert (Bethke B and Sauer B (1997) Nucl Acids Res 25:2828-2834).
- 10 Die "Replacement" Strategie wurde auch bereits im Kern von Schizisacchomyces pombe mit der Rekombinase Φ C31 verwirklicht (Thomason LC (2001) Mol Gen Genet 265:1031-1038). Beim Nutzen der "Replacement" Strategie wird ein eventuell mit dem RE-Konstrukt
- 15 eingebrachter Selektionsmarker bevorzugt zwischen den Erkennungsregionen R1 und R1' positioniert, so dass er bei Insertion der Insertionssequenz flankiert von R2 und R2' direkt aus der Masterpflanze eliminiert wird.

- Bevorzugt können bei der "Replacement" Strategie 2 RE-Sequenzen
- 20 (R1 und R1' bzw. R2 und R2') (beispielsweise jeweils 2 lox Sequenzen) genutzt werden, die nicht miteinander rekombinieren können, weil sie sich beispielsweise im Spacer unterscheiden (siehe auch Bethke B et al. (1997) Nucl Acids Res 25:2828-2834). Ähnliches auch für das FLP/frt beschrieben worden, bei dem ins-
- 25 besondere die meisten Nukleotide der Spacer Region ausgetauscht werden können (Schlake, T & Bode, J, 1994, Biochemistry 33: 12746-12751).

- Für die Replacement Strategie können auch 2 gleiche attB bzw.
- 30 attP Sites genutzt werden. Diese können nicht untereinander rekombinieren.

- In einer weiteren - besonders bevorzugten Ausführungsform - enthält das RE-Konstrukt bevorzugt zusätzlich zur Erkennungsstelle
- 35 für eine Rekombinase auch mindestens eine Erkennungsstelle für ein DSBI-Enzym. Diese Konstrukte werden infolge auch als RE/DSB-Konstrukt bezeichnet, sind aber auch unter dem allgemeinen Begriff des RE-Konstruktes subsumiert.

- 40 Das jeweilige DSBI-Enzym besitzt - außer der über das RE/DSB-Konstrukt eingebrachten - bevorzugt keine weiteren DSB-Erkennungsstellen im Plasmid (außer: Duplikationen beispielsweise in "repeats"). Es ist nicht zwingend notwendig, dass die DSB-Erkennungssequenz auf dem RE/DSB-Konstrukt vorhanden ist.
- 45 Sie kann auch natürlicherweise im Plasmid benachbart zu der Insertionsstelle eines RE-Konstruktes vorhanden sein. Eine funktionelle RE/DSB-Einheit entsteht dann erst nach der Inte-

47

gration des RE-Konstruktes in das Plastom. Mit Hilfe des RE/DSB-Konstruktes wird zunächst eine überwiegend homoplastome Masterpflanze erzeugt, wobei dem Fachmann vertraute Verfahren wie beispielsweise die Homologe Rekombination eingesetzt werden.

5

In einer weiteren Ausführungsform wird durch die Insertionssequenz zusätzlich eine Sequenz X in das Plastom eingeführt, welche homolog zu einer bereits im Plastom vorhandenen Sequenz X' der Masterpflanze ist. Diese Sequenz X' kann natürlicherweise vor-

10 handen sein, oder aber durch das RE/DSB-Konstrukt eingeführt werden. Die Sequenzen X und X' sind so lokalisiert, dass zwischen ihnen - nach der Integration der Insertionssequenz mittels Rekombination - nachfolgende Elemente zu liegen kommen:

15 i) eine bevorzugt singuläre DSB-Erkennungssequenz für ein DSBI-Enzym

ii) hybride RE-Sequenz R1/2 der Rekombinase, die sich nach dem Einbau des sekundären Transformationsvektors durch Rekombi-
20 nation von R1 mit R2 ergeben hat

iii) ggf. weitere Elemente wie beispielsweise positive oder negative Selektionsmarker

25 wobei X und X' eine Orientierung sowie ausreichende Homologie und Länge besitzen, um infolge eines sequenzspezifischen Doppelstrangbruches an der DSB-Erkennungssequenz homolog mit einander zu rekombinieren und eine Deletion der zwischen ihnen gelegenen Sequenzen zu bewirken.

30

Das Transformationskonstrukt, der Transformationsvektor oder die Insertionssequenz werden wie oben beschrieben in die Plastiden eingebracht. Durch die Aktivität einer entsprechenden Rekombinase wird die Insertionssequenz in das Plastom eingebaut. Bevorzugt

35 wird ein System genutzt, welches in der eingesetzten Form nahezu oder vollständig irreversibel ist. Besonders bevorzugt werden die Rekombinasen der Phagen Φ C31 oder TP901-1 und ihre entsprechenden Erkennungsregionen genutzt.

40 Nach Integration der Insertionssequenz wird in einer besonders bevorzugten Ausführungsform eine entsprechendes DSBI-Enzym induziert, das an seiner DSB-Erkennungssequenz im Plastom der Masterpflanze ein DNA Doppelstrangbruch erzeugt.

45

Nun sind zwei Fälle zu unterscheiden:

- 5 a) Die betrachtete Kopie der plastidären DNA ist bereits in der gewünschten Weise durch Rekombinase vermittelte Insertion der Insertionssequenz modifiziert worden. Der DNA Doppelstrangbruch erfolgt an einer Erkennungsregion, die benachbart zu besagter Insertionssequenz lokalisiert ist. Durch intramolekulare homologe Rekombination zwischen X und X' wird nun die zwischen diesen lokalisierte DSB-Erkennungsstelle entfernt oder zumindest derart zerstört werden, dass sie nicht mehr 10 von dem DSBI-Enzym erkannt wird. In Folge können die so modifizierten plastidären DNA-Moleküle nicht mehr durch das DSBI-Enzym geschnitten werden. Sie können störungsfrei replizieren und/oder als Substrate für Reperatur-synthesen weiterer - noch nicht modifizierter - plastidärer DNA-Moleküle fungieren. In 15 einer Variante dieser Ausführungsform wird durch die induzierte Deletion ein negativer Selektionsmarker deletiert, was eine erleichterte Selektion der plastidären DNA-Moleküle mit Deletion ermöglicht.
- 20 b) Die betrachtete Kopie plastidäre DNA ist noch nicht in der gewünschten Weise durch Rekombinase vermittelte Insertion der Insertionssequenz modifiziert worden. Auch in diesem Fall ergibt sich ein DNA-Doppelstrangbruch. Hier steht jedoch 25 intramolekular keine homologe Region zur Verfügung, die zur Reparatur des Doppelstrangbruches genutzt werden könnte. Es ist intramolekular keine Rekombination zwischen X und X' möglich, demnach auch keine Deletion der DSB-Erkennungssequenz. Die so geschnittenen plastidären DNA-Moleküle können zum 30 einen nicht mehr ungestört replizieren. Zum anderen können besagte geschnittene Moleküle im Rahmen einer Reperatur-synthese durch ein Genkonversionereignis unter Verwendung von Kopien der bereits in der gewünschten Weise modifizierten, nicht mehr schneidbaren plastidären DNA-Moleküle (die 35 also in Folge der oben unter a) beschriebenen Ereignisse keine DSB-Erkennungssequenz mehr enthalten) repariert werden. So können auch ohne selektive Regeneration das Transformationskonstrukt / die Insertionssequenz in alle oder zumindest die überwiegende Anzahl der Plastomkopien eines 40 Plastids verteilt werden. Dies Verfahren beschleunigt den Prozess der Plastidentransformation und verringert aufgrund weniger notwendiger Regenerationsschritte die Gefahr der Anhäufung von somaklonalen Variationen. Außerdem ermöglicht es, eine Verteilung des Transgens in die Kopien des Plastoms zu erzielen, ohne einen Selektionsmarker und entsprechendes 45 selektives Agens nutzen zu müssen.

Das Transformationskonstrukt kann weitere Elemente umfassen. So kann es wünschenswert sein, zwischen Sequenzbereich X und der DSB-Erkennungssequenz mindestens ein Gen für einen positiven und/oder negativen Selektionsmarker zu platzieren. Durch die induzierte Rekombination werden so auch die Selektionsmarker bzw. jeder andere beliebige Sequenzabschnitt, der sich zwischen den homologen Bereichen X und X' befindet, eliminiert. Entsprechend kann auf diese Weise beispielsweise das Vektorrückgrat - wenn es nicht vor der Insertion entfernt wurde - wieder aus dem Plasmid eliminiert werden.

Als DSBE-Enzym wird bevorzugt die Endonuklease I-PpoI aus *Physarum polycephalum* genutzt. Es ist aber auch möglich, jedes andere DSBE-Enzym zu nutzen, das in den betrachteten Organell aktiv ist.

Das DSBE-Enzym wird bevorzugt erst nach der Rekombinase, die für den Einbau des Transformationsvektors / der Insertionssequenz verantwortlich ist, aktiviert. Auch hier sind verschiedene Möglichkeiten denkbar, dies zu realisieren. Entsprechende Verfahren sind bereits für die Rekombinase beschrieben worden und sind auf das DSBE-Enzym übertragbar. So kann das DSBE-Enzym in Fusion mit einem Transitpeptid im Kern der Masterpflanze exprimiert werden. Die Expression sollte bevorzugt induzierbar sein oder im Fall von konstitutiver Expression sollte die Aktivität der entsprechenden Endonuklease aktivierbar sein. Darüber hinaus kann das DSBE-Enzym als Fusion mit einem Transitpeptid in Agrobakterien exprimiert und als Fusionsprotein beispielsweise mit VirF in Pflanzenzellen eingebracht werden. Auch transiente Expression vom Kern mit post-translationalem Import in die betrachteten Plastiden ist denkbar. Bevorzugt wird das DSBE-Enzym bzw. eine das DSBE-Enzym kodierende Nukleinsäuresequenz jedoch erst mit dem Transformationsvektor in die Plastiden der Masterpflanze eingebracht. Dies kann dadurch geschehen, daß das DSBE-Enzym als Protein gleichzeitig mit dem Transformationskonstrukt/ der Insertionssequenz in die Organellen eingebracht wird; daß eine in vitro erstellte RNA, die das DSBE-Enzym kodiert, mit dem Transformationskonstrukt/ der Insertionssequenz in die Plastiden eingebracht wird; oder dass - besonders bevorzugt - eine das DSBE-Enzym kodierende DNA mit dem Transformationskonstrukt/ der Insertionssequenz in die Plastiden eingebracht wird. Ganz besonders bevorzugt ist das DSBE-Enzym in letzterem Fall in einer in dem Plastid exprimierbaren Form codiert. Besonders bevorzugt ist das DSBE-Enzym auf der Insertionssequenz codiert. Ganz besonders bevorzugt wird die Expression des DSBE-Enzyms erst nach der Insertion der Insertionssequenz ermöglicht. Dies kann beispielsweise dadurch erreicht werden, daß auf dem Transformationskonstrukt / der Insertionssequenz keine regulatorischen Elemente vorhanden sind, die die Expression des DSBE-Enzymes steuern. Die Expression des DSBE-Enzymes könnte in einem

50

- solchen Fall beispielsweise nach Integration in die plastidäre DNA auch durch regulatorische Elemente aus dem Plastom der Masterpflanze gesteuert werden. Falls die Expressionskassette für das DSBI-Enzym auf der Insertionssequenz enthalten ist, wird
- 5 - besonders bevorzugt - das DSBI-Enzym von Bereichen flankiert, die eine spätere Excision des DSBI-Enzyms erlauben.

- In einer bevorzugten Ausführungsform umfasst die Insertionssequenz mindestens eine weitere zu exprimierende Nukleinsäuresequenz, die bevorzugt einen gewünschten Phänotyp in einer entsprechend transformierten Pflanze erzeugt. Um die Expression (Transkription und/oder Translation) zu gewährleisten, sind diese - je nach Ausführungsform und Insertionsort - mit regulatorischen Elementen zu versehen. Erfolgt die Insertion in einen trans-
- 15 kriptionell aktiven Locus so sind - wie oben beschrieben - keine Promotorsequenzen erforderlich. Vorteilhaft werden die zu exprimierenden Sequenzen in jedem Fall mit Ribosomenbindestellen in geeignetem Abstand stromaufwärts des offenen Leserahmens versehen oder besitzen bereits solche von Natur aus. Diese regulatorischen
- 20 Sequenzen oder Teile derselben können aber auch im Plastom natürlicherweise vorhanden sein oder bereits im ersten Schritt, d.h. bei der Generierung einer nicht-natürlichen Masterpflanze zusammen mit der der RE-Sequenz in die plastidäre-DNA eingeführt werden.

- 25 "Ausreichende Länge" meint in Bezug auf die Homologiesequenzen (z.B. A, A', B, B', X, X', H1, H2) bevorzugt Sequenzen von einer Länge von mindestens 20 Basenpaaren, bevorzugt mindestens 50 Basenpaaren, besonders bevorzugt von mindestens 100 Basen-
- 30 paaren, ganz besonders bevorzugt von mindestens 250 Basenpaaren, am meistens bevorzugt von mindestens 500 Basenpaaren.

- "Ausreichende Homologie" meint in Bezug auf die Homologiesequenzen (z.B. A, A', B, B', X, X', H1, H2) bevorzugt Sequenzen
- 35 die eine Homologie innerhalb dieser Homologiesequenzen aufweisen von mindestens 70 %, bevorzugt 80 %, vorzugsweise mindestens 90 %, besonders bevorzugt mindestens 95 %, ganz besonders bevorzugt mindestens 99 %, am meisten bevorzugt 100 % über eine Länge von mindestens 20 Basenpaaren, bevorzugt mindestens 50 Basen-
- 40 paaren, besonders bevorzugt von mindestens 100 Basenpaaren, ganz besonders bevorzugt von mindestens 250 Basenpaaren, am meistens bevorzugt von mindestens 500 Basenpaaren.

- Unter Homologie zwischen zwei Nukleinsäuren wird die Identität
- 45 der Nukleinsäuresequenz über die jeweils gesamte Sequenzlänge verstanden, die durch Vergleich mit Hilfe des Programmalgorithmus GAP (Wisconsin Package Version 10.0, University of Wisconsin,

51

Genetics Computer Group (GCG), Madison, USA) unter Einstellung folgender Parameter berechnet wird:

Gap Weight: 12

Length Weight: 4

5

Average Match: 2,912

Average Mismatch:-2,003

Bevorzugt umfasst das Transformationskonstrukt oder die Insertionssequenz einen Selektionsmarker, der eine Selektion transplastomer Plastiden ermöglicht (s.u.), besonders bevorzugt aadA, BADH oder einen "binding type"-Marker. Der Selektionsmarker ist bevorzugt so konstruiert, dass eine nachfolgende Deletion aus dem Plastom ermöglicht wird. Entsprechende Verfahren sind dem Fachmann bekannt und unten beschrieben.

15

Die Insertionssequenz, das Transformationskonstrukt, das RE-Konstrukt oder die Expressionskassetten für eine Rekombinase oder ein DSBI-Enzym können zum Aufbau eines Transformationsvektor in einen Standard-Vektor wie pBluescript oder pUC18 kloniert werden.

20 In einer bevorzugten Ausführungsform wird die Insertionssequenz bzw. Transformationskonstrukt als lineares oder linearisiertes DNA-Molekül appliziert.

Bevorzugt wird nur der Teil des Transformationsvektors appliziert, der die Insertionssequenz oder das Transformationskonstrukt mit RE-Sequenz, Selektionsmarker und/oder die Expressionskassette für das DSBI-Enzym umfasst.

Eines der oben beschriebenen Konstrukte kann mit einem der beschriebenen Verfahren in die Plastiden einer entsprechenden Masterpflanze eingebracht werden. Bevorzugt sind Mikroinjektion und besonders bevorzugt Partikelbeschuss.

Klonierungs-, Expressions-, Selektions- und Transformationsverfahren

"Expressionskassette" meint - zum Beispiel in Bezug auf die Expressionkassette für eine Rekombinase oder ein DSBI-Enzym - solche Konstruktionen bei denen die zu exprimierende DNA in funktioneller Verknüpfung mit mindestens einem genetischen Kontrollelement steht, dass ihre Expression (d.h. Transkription und oder Translation) ermöglicht oder reguliert. Dabei kann die Expression zum Beispiel stabil oder transient, konstitutiv oder induzierbar erfolgen. Für die Einführung stehen dem Fachmann verschiedene unten aufgeführte direkte (z.B. Transfektion, Partikelbeschuss, Mikroinjektion) oder indirekte Verfahren (z.B. Agrobakterien-

infektion, Virusinfektion) zur Verfügung, die weiter unten aufgeführt werden.

Unter einer funktionellen Verknüpfung versteht man allgemein eine
5 Anordnung in der eine genetische Kontrollsequenz ihre Funktion in
Bezug auf eine Nukleinsäuresequenz - beispielsweise kodierend für
eine Rekombinase oder ein DSBI-Enzym - ausüben kann. Funktion
kann dabei beispielsweise die Kontrolle der Expression d.h. Tran-
skription und/oder Translation der Nukleinsäuresequenz - bei-
10 spielsweise kodierend für eine Rekombinase oder ein DSBI-Enzym -
bedeuten. Kontrolle umfasst dabei beispielsweise die Initiierung,
Steigerung, Steuerung oder Suppression der Expression d.h. Tran-
skription und ggf. Translation. Die Steuerung wiederum kann bei-
spielsweise gewebe- und oder zeitspezifisch erfolgen. Sie kann
15 auch induzierbar zum Beispiel durch bestimmte Chemikalien,
Stress, Pathogene etc. sein.

Unter einer funktionellen Verknüpfung versteht man zum Beispiel
die sequentielle Anordnung eines Promoters, der zu exprimierenden
20 Nukleinsäuresequenz - beispielsweise kodierend für eine Rekombi-
nase oder ein DSBI-Enzym - und ggf. weiterer regulativer Elemente
wie zum Beispiel einem Terminator derart, dass jedes der regu-
lativen Elemente seine Funktion bei der Expression der Nuklein-
säuresequenz - beispielsweise kodierend für eine Rekombinase oder
25 ein DSBI-Enzym - erfüllen kann. Dabei ist es nicht zwingend not-
wendig, dass die funktionelle Verknüpfung bereits auf den Trans-
formationskonstrukten gegeben ist. Die funktionelle Verknüpfung
kann sich auch in Folge der Insertion in die Kern- bzw. plasti-
däre DNA ergeben, wobei hier die regulativen Elemente bereits
30 in der Kern- bzw. plastidären DNA vorliegen. Die regulativen
Elemente können diesbezüglich natürlicherweise vorliegen oder
aber in einem vorgeschalteten Schritt - beispielsweise bei der
Einführung einer RE-Sequenz - eingebracht werden.

35 Dazu ist nicht unbedingt eine direkte Verknüpfung im chemischen
Sinne erforderlich. Genetische Kontrollsequenzen, wie zum Bei-
spiel Enhancer-Sequenzen, können ihre Funktion auch von weiter
entfernten Positionen oder gar von anderen DNA-Molekülen aus auf
die Zielsequenz ausüben. Bevorzugt sind Anordnungen, in denen die
40 zu exprimierende Nukleinsäuresequenz - beispielsweise codierend
für eine Rekombinase oder ein DSBI-Enzym - hinter eine als
Promoter fungierende Sequenz positioniert wird, so dass beide
Sequenzen kovalent miteinander verbunden sind. Bevorzugt ist
dabei der Abstand zwischen der Promotorsequenz und der Nuklein-
45 säuresequenz - beispielsweise kodierend für eine Rekombinase oder
ein DSBI-Enzym - geringer als 200 Basenpaare, besonders bevorzugt

kleiner als 100 Basenpaare, ganz besonders bevorzugt kleiner als 50 Basenpaare.

- Dem Fachmann sind verschiedene Wege bekannt, um zu einem
- 5 der erfindungsgemäßen Transformationskonstrukte, diese umfassenden Vektoren oder einer der Expressionskassetten zu gelangen. Die Herstellung kann mittels gängiger Rekombinations- und Klonierungstechniken realisiert werden, wie sie beispielsweise in Maniatis T, Fritsch EF und Sambrook J, Molecular Cloning: A
- 10 Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1989) sowie in Silhavy TJ, Berman ML und Enquist LW, Experiments with Gene Fusions, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1984) und in Ausubel FM et al., Current Protocols in Molecular Biology, Greene Publishing
- 15 Assoc. and Wiley Interscience (1987) beschrieben sind. Bevorzugt ist die direkte Fusion einer als Promoter fungierenden Nukleinsäuresequenz mit einer zu exprimierenden Nukleotidsequenz - beispielsweise kodierend für eine Rekombinase oder ein DSBI-Enzym.
- 20 Der Begriff der "genetischen Kontrollsequenzen" ist breit zu verstehen und meint all solche Sequenzen, die einen Einfluss auf das Zustandekommen oder die Funktion einer Expressionskassette oder Transformationsvektors haben. Genetische Kontrollsequenzen gewährleisten die Transkription und gegebenenfalls
- 25 Translation in Zellkern (bzw. Zytoplasma) oder Plastiden. Vorzugsweise umfassen die erfindungsgemäßen Expressionskassetten 5'-stromaufwärts von der jeweiligen zu exprimierenden Nukleinsäuresequenz einen Promotor und 3'-stromabwärts eine Terminatorsequenz als zusätzliche genetische Kontrollsequenz, sowie
- 30 gegebenenfalls weitere übliche regulative Elemente, und zwar jeweils funktionell verknüpft mit der zu exprimierenden Nukleinsäuresequenz.

- Genetische Kontrollsequenzen sind beispielsweise beschrieben bei
- 35 "Goeddel; Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, CA (1990)" oder "Gruber and Crosby, in: Methods in Plant Molecular Biology and Biotechnology, CRC Press, Boca Raton, Florida, eds.: Glick and Thompson, Chapter 7, 89-108" sowie den dort aufgewiesenen Zitaten.

- 40 Beispiele für derartige Kontrollsequenzen sind Sequenzen, an die Induktoren oder Repressoren binden und so die Expression der Nukleinsäure regulieren. Zusätzlich zu diesen neuen Kontrollsequenzen oder anstelle dieser Sequenzen kann die natürliche
- 45 Regulation dieser Sequenzen vor den eigentlichen Strukturgenen noch vorhanden sein und gegebenenfalls genetisch verändert worden sein, so dass die natürliche Regulation ausgeschaltet und die

54

Expression der Gene erhöht wurde. Die Expressionskassette kann aber auch einfacher aufgebaut sein, das heißt, es werden keine zusätzlichen Regulationssignale vor die vorstehend erwähnten Gene inseriert und der natürliche Promotor mit seiner Regulation wird
5 nicht entfernt. Stattdessen wird die natürliche Kontrollsequenz so mutiert, dass keine Regulation mehr erfolgt und die Genexpression gesteigert wird. Diese veränderten Promotoren können auch allein vor die natürlichen Gene zur Steigerung der Aktivität gebracht werden.

10

Je nach vorstehend näher beschriebenen pflanzlichen Organismen, die als Wirts- oder Ausgangsorganismus zum Einsatz kommen können, und der durch Einbringen der Expressionskassetten oder Vektoren in einen genetisch veränderten oder transgenen Organismus über-
15 führt wird, eignen sich verschiedene Kontrollsequenzen.

Zur nukleären Expression (beispielsweise einer viralen/Bakteriophagen RNA-Polymerase, einer Rekombinase oder eines DSBI-Enzyms jeweils als Fusionsprotein mit einem plastidären Transitpeptid)
20 sind grundsätzlich alle Promotoren geeignet, die die Expression von Genen, insbesondere Fremdgenen, in Pflanzen steuern können.

Geeignet sind Promotoren, die eine konstitutive Expression in Pflanzen ermöglichen (Benfey et al. (1989) EMBO J 8:2195-2202).
25 Vorzugsweise verwendet man insbesondere einen pflanzlichen Promotor oder einen Promotor, der einem Pflanzenvirus entstammt. Besonders bevorzugt ist der Promotor des 35S-Transkriptes des Blumenkohlmosaikvirus (Franck et al. (1980) Cell 21:285-294; Odell et al. (1985) Nature 313:810-812; Shewmaker et al. (1985) Virology 140:281-288; Gardner et al. 1986, Plant Mol. Biol. 6, 221-228) oder den 19S CaMV Promotor (US 5,352,605 and WO 84/02913). Dieser Promotor enthält bekanntlich unterschiedliche Erkennungssequenzen für transkriptionale Effektoren, die in ihrer Gesamtheit zu einer weitgehend permanenten und konsti-
30 tuten Expression des eingeführten Gens führen (Benfey et al. (1989) EMBO J 8:2195-2202). Ein weiterer geeigneter konstitutiver Promotor ist der "Rubisco small subunit (SSU)"-Promotor (US 4,962,028). Ein weiteres Beispiel eines geeigneten Promotors ist der LeguminB-Promotor (GenBank Acc.-No.: X03677). Weitere bevor-
40 zugte konstitutive Promotoren sind zum Beispiel der Promotor der Nopalinsynthase aus Agrobakterium, der TR-Doppelpromotor, der OCS (Octopin Synthase) Promotor aus Agrobakterium, der Ubiquitin Promotor (Holtorf S et al. (1995) Plant Mol Biol 29:637-649), die Promotoren der vakuolärer ATPase Untereinheiten, der FBPaseP
45 Promotor (WO 98/18940) oder der Promotor eines prolinreichen Proteins aus Weizen (WO 91/13991). Weiterhin geeignete und im Rahmen dieser Erfindung bevorzugte konstitutive Promotoren sind

- der SuperPromotor (Ni M et al. (1995) Plant J 7:661-676; US 5,955,646) sowie der Nitrilase-1 Promotor des nit1 Gens aus Arabidopsis (GenBank Acc.-No.: Y07648.2, Nukleotide 2456 bis 4340; Hillebrand H et al. (1998) Plant Mol Biol 36 (1):89-99;
- 5 Hillebrand H et al. (1996) Gene 170(2):197-200).

- Bevorzugt sind induzierbare, besonders bevorzugt chemisch induzierbare Promotoren (Aoyama T und Chua NH (1997) Plant J 11:605-612; Caddick MX et al. (1998) Nat. Biotechnol 16:177-180;
- 10 Rewiew: Gatz (1997) Annu Rev Plant Physiol Plant Mol Biol 48:89-108), durch die die Expression zu einem bestimmten Zeitpunkt gesteuert werden kann. Beispielhaft seien zu nennen der PRP1 Promotor (Ward et al. (1993) Plant Mol Biol 22:361-366), ein durch Salicylsäure induzierbarer Promotor (WO 95/19443),
- 15 ein durch Benzolsulfonamid-induzierbarer Promotor (EP-A-0388186), ein durch Tetrazyklin-induzierbarer Promotor (Gatz et al. (1992) Plant J 2:397-404), ein durch Abscisinsäure-induzierbarer Promotor (EP-A 335 528), ein durch Ethanol induzierbarer Promotor (Salter MG et al. (1998) Plant J. 16:127-132), der schwermetall-
- 20 induzierbarer Metallothionein I Promotor (Amini S et al. (1986) Mol Cell Biol 6:2305-2316), der steroid-induzierbarer MMTV LTR Promotor (Izant JG et al. (1985) Science 229:345-352) sowie ein durch Cyclohexanon-induzierbarer Promotor (WO 93/21334). Besonders bevorzugt ist die induzierbare Expression eines PLS/DSBI-Enzym-Fusionsproteins im Kern. Induzierbare Promotoren umfasst auch solche, die durch bestimmte Repressorproteine (z.B. tet, lac) reguliert werden können. Entsprechende Repressorproteine können in Fusion mit PLS in die Plastiden translokalisiert werden und dort die Expression bestimmter Gene unter Kontrolle entsprechender
- 30 Promotoren regulieren. Der Repressor bindet in den Plastiden an eine künstliche ins Plastom eingefügte Repressorbindestelle und kann so die Expression des stromabwärts gelegenen Gens unterdrücken (vgl. WO95/25787). Dadurch kann beispielsweise die Expression einer plastidkodierten Rekombinase oder eines DSBI-Enzym im
- 35 Bedarfsfall induziert oder sie bis auf den Moment, in dem die Expression gewünscht ist, unterdrückt werden.

- Ferner sind Promotoren bevorzugt, die durch biotischen oder abiotischen Stress induziert werden wie beispielsweise der
- 40 pathogen-induzierbare Promotor des PRP1-Gens (Ward et al., Plant Mol Biol 1993, 22: 361-366), der hitzeinduzierbare hsp70-Promotor oder der hsp80-Promoter aus Tomate (US 5,187,267), der kälteinduzierbare alpha-Amylase Promoter aus der Kartoffel (WO 96/12814) oder der verwundungsinduzierte pinII-Promoter
- 45 (EP-A 0 375 091).

56

In einer besonders bevorzugten Ausführungsform wird vor allem die für die Rekombinase oder das DSBI-Enzym kodierende Nukleinsäure unter Kontrolle eines induzierbaren Promotors exprimiert. Damit wird eine kontrollierte, steuerbare Expression erreicht
5 und etwaige Probleme durch eine konstitutive Expression einer Rekombinase oder eines DSBI-Enzyms vermieden.

Vorteilhafte Kontrollsequenzen für die erfindungsgemäßen Expressionskassetten oder Vektoren umfassen virale, bakteriophage
10 oder bakterielle Promotoren wie cos-, tac-, trp-, tet-, phoA-, tat-, lpp-, lac-, lacIq-, T7-, T5-, T3-, gal-, trc-, ara-, SP6-, λ -PR- oder im λ -PL-Promotor. Diese werden bevorzugt in Kombination mit der Expression der jeweils korrespondierenden RNA-Polymerase eingesetzt.

15

Die Expression in Plastiden kann unter Verwendung plastidärer Promotoren und/oder Transkriptionsregulationselemente realisiert werden. Beispielfhaft jedoch nicht einschränkend seien zu nennen der RNA-Polymerase Promotor (WO 97/06250) oder die in
20 WO 00/07431, US 5,877,402, WO 97/06250, WO 98/5559,5 WO 99/46394, WO 01/42441 und WO 01/07590 beschriebenen Promotoren. Zu nennen sind das rpo B Promotorelement, das atpB Promotorelement, das clpP Promotorelement (siehe auch WO 99/46394) oder das 16S-rDNA Promotorelement. Dabei kann dem Promotor auch ein polycistroni-
25 sches "Operon" zugeordnet sein (EP-A 1 076 095; WO 00/20611). Beschrieben sind auch Systeme bei denen eine nicht-pflanzliche (z.B. virale) RNA-Polymerase unter Verwendung plastidärer Trans-
itpeptide in das Plastid importiert wird und dort spezifisch die Expression transgener Sequenzen induziert, die unter Kontrolle
30 der Erkennungssequenzen der RNA-Polymerase stehen und zuvor in die plastidäre DNA insertiert wurden (WO 95/16783; US 5,925,806; US 5,575,198).

Neben den genannten Promotoren wären weiter bevorzugt zu ver-
35 wenden:

a) der PrbcL Promotor (SEQ ID NO: 20)

b) der Prps16 Promotor (SEQ ID NO: 26)

40

c) der Prrn16 Promotor (SEQ ID NO: 22)

In einer bevorzugten Ausführungsform werden NEP Promotoren eingesetzt. Dies sind Promotoren, die in Plastiden funktionell sind
45 und von der nukleär kodierten, plastidären RNA-Polymerasen (NEP) erkannt werden. Bevorzugt sind: Prrn-62; Pycf2-1577; PatpB-289; Prps2-152; Prps16-107; Pycf1-41; PatpI-207; PclpP-511; PclpP-173

und PaccD-129 (WO 97/06250; Hajdukiewicz PTJ et al (1997) EMBO J 16:4041-4048).

Besonders bevorzugt sind:

5

a) Der PaccD-129 Promotor des accD Gens aus Tabak (WO 97/06250; SEQ ID NO: 23)

10 b) Der PclpP-53 Promotor des clpP Gens als hoch-aktiver NEP Promotor in Chloroplasten (WO 97/06250; SEQ ID NO: 24)

c) Der Prrn-62 Promotor des rrn Gens (SEQ ID NO: 25)

15 d) Der Prps16-107 Promotor des rps16 Gens (SEQ ID NO: 21):

e) Der PatpB/E-290 Promotor des atpB/E Gens aus Tabak (Kapoor S et al. (1997) Plant J 11:327-337) (SEQ ID NO: 27)

20 f) Der PrpoB-345 Promotor des rpoB Gens (Liere K & Maliga P (1999) EMBO J 18: 249-257) (SEQ ID NO: 28)

Im Allgemeinen sind in dieser Ausführungsform alle Promotoren nutzbar, die der Klasse III angehören (Hajdukiewicz PTJ et al (1997) EMBO J 16:4041-4048) sowie alle Fragmente der Promotoren
25 der Klasse II, die die Transkriptionsinitiation durch die NEP steuern. Solche Promotoren bzw. Promotor-Anteile sind nicht besonders hoch konserviert. Als Consensus nahe der Transkriptionsinitiationsstelle von NEP Promotoren wird angegeben: ATAGAATAAA (Hajdukiewicz PTJ et al (1997) EMBO J 16:4041-4048).

30

Normalerweise werden Gene von regulatorischen Sequenzen umgeben, die aus den Plastiden des zu transformierenden Organismus stammen. Dadurch erzeugt man Sequenzduplikationen, die zu Instabilitäten aufgrund von spontanen, intrachromosomalen homologen

35 Rekombinationsereignissen führen können (Heifetz PB (2000) Biochimie 82(6-7):655-666). Es wurde zur Lösung dieses Problems vorgeschlagen, heterologe regulatorische Sequenzen zu nutzen, oder endogen im Plastidengenom bereits vorhandene regulatorische Einheiten auszunutzen (WO 99/46394; WO 01/42441). Auch eine Ver-
40 minderung der Homologie durch Mutagenese der endogenen Promotorsequenz ist beschrieben (WO 01/07590).

Prinzipiell können alle natürlichen Promotoren mit ihren Regulationssequenzen wie die oben genannten für das erfindungs-
45 gemäße Verfahren verwendet werden. Besonders bevorzugt sind Promotoren isoliert aus Prokaryoten. Ganz besonders bevorzugt sind Promotoren isoliert aus Synechocystis oder E.coli. Darüber

hinaus können auch synthetische Promotoren vorteilhaft verwendet werden, wie beispielsweise ein synthetischer Promotor abgeleitet von der E. coli Konsensus Sequenz für $\sigma 70$ Promotoren:

5 5-TTGACA N₁₆₋₁₉ TATAAT N₃ CAT -3',

wobei N für jedes beliebige Nukleotid (also A, G, C oder T) steht. Dem Fachmann ist offensichtlich, dass auch einzelne bis wenige Basenaustausche in den angegebenen, konservierten Regionen
 10 möglich sind, ohne die Funktion des Promotors zu zerstören. Die variable Gestaltung dieser synthetischen Promotoren durch Nutzen verschiedener Sequenzabfolgen, ermöglicht es, eine Vielzahl von Promotoren zu erstellen, die nicht extensive Homologien aufweisen, was insbesondere für den Fall, dass mehrere Promotoren
 15 benötigt werden, die Stabilität der Expressionskassetten im Plasmid erhöht. Beispielfhaft aber nicht einschränkend seien nachfolgende besonders bevorzugte Promotorsequenzen genannt, die von oben genannter Konsensussequenz abgeleitet sind:

- 20 a) 5'-TTGACATTCACCTCTTCAATTATCTATAATGATACA-3' (SEQ ID NO: 29)
 b) 5'-TTGACAATTTTCCTCTGAATTATATAATTAACAT-3' (SEQ ID NO: 59)

Dem Fachmann ist offensichtlich, dass diese synthetischen Promotoren die Expression beliebiger Gene steuern können. Sie
 25 können beispielsweise dazu genutzt werden, die Expression eines Selektionsmarkers anzutreiben - auch, um unter regenerativen Bedingungen auf transplastome Pflanzen unter Zuhilfenahme von besagtem Selektionssystem selektieren zu können. Selektionsmarker sind weiter unten beispielhaft aufgezählt. Darüber hinaus können
 30 solche synthetischen Promotoren mit jedem beliebigen Gen verknüpft werden, beispielsweise mit Genen kodierend für Antikörper, Antigene oder Enzyme. Bevorzugt enthalten die Expressionskassetten bestehend aus solchen Promotoren auch im folgenden näher beschriebene 5'-untranslatierte Regionen (oder Ribosomen-
 35 bindestellen) oder 3'-nichtkodierende Regionen.

Genetische Kontrollsequenzen umfassen auch weitere Promotoren, Promotorelemente oder Minimalpromotoren, die die expressionssteuernden Eigenschaften modifizieren können. Genetische Kon-
 40 trollsequenzen umfassen ferner auch die 5'-untranslatierte Region (5'-UTR) oder die nichtkodierende 3'-Region (3'-UTR) von Genen (Eibl C (1999) Plant J 19:1-13). Es ist gezeigt worden, dass diese eine signifikante Funktionen bei der Regulation der Genexpression in Plastiden höherer Pflanzen spielen können. Auch im
 45 Kern können genetische Kontrollelemente wie 5'-UTR, Introns oder 3'UTR eine Funktion bei der Genexpression besitzen. So wurde beispielsweise gezeigt, dass 5'-untranslatierte Sequenzen die trans-

59

iente Expression heterologer Gene verstärken können. Sie können ferner die Gewebespezifität fördern (Rouster J et al. (1998) Plant J 15:435-440).

5 Bevorzugt werden als 5' UTRs und 3'UTRs eingesetzt:

- a) 5'psbA (aus Tabak) (SEQ ID NO: 30)
- b) 5'rbcl einschließlich 5' Anteilen aus dem codierenden Bereich des rbcl Gens (aus Tabak) (SEQ ID NO: 31); die mit SEQ ID NO: 31 beschriebene Sequenz wurden gegenüber der nativen Sequenz mutiert, um eine PstI bzw. NcoI Schnittstelle einzubringen.
- 10 c) 5'rbcl (SEQ ID NO: 32); die mit SEQ ID NO: 32 beschriebene Sequenz wurden gegenüber der nativen Sequenz mutiert, um eine PstI Schnittstelle einzubringen.
- 15 c) 5'rbcl (SEQ ID NO: 32); die mit SEQ ID NO: 32 beschriebene Sequenz wurden gegenüber der nativen Sequenz mutiert, um eine PstI Schnittstelle einzubringen.
- d) 3'psbA-1 aus Synechocystis (SEQ ID NO: 33)
- 20 e) 3'psbA aus Tabak (SEQ ID NO: 34)
- f) 3'rbcl aus Tabak (SEQ ID NO: 35)
- 25 Genetische Kontrollsequenzen vor allem für die Expression in Plastiden umfassen insbesondere auch Ribosomenbindungssequenzen zur Initiation der Translation. Diese sind für gewöhnlich in den 5'UTRs enthalten. Dies ist vor allem dann bevorzugt, wenn von der zu exprimierenden Nukleinsäuresequenz entsprechende
- 30 Sequenzen nicht bereitgestellt werden oder diese mit dem Expressionssystem nicht kompatibel sind. Insbesondere bevorzugt ist die Verwendung einer synthetische Ribosomenbindestelle (RBS) mit der Sequenz 5'-GGAGG(N)₃₋₁₀ATG-3', bevorzugt 5'-GGAGG(N)₅ATG-3' (SEQ ID NO: 36) besonders bevorzugt 5'-GGAG-
- 35 GATCTCATG-3' (SEQ ID NO: 37).

Die Expressionskassette kann vorteilhafterweise eine oder mehrere sogenannte "enhancer Sequenzen" funktionell verknüpft mit dem Promoter enthalten, die eine erhöhte transgene Expression der

- 40 Nukleinsäuresequenz ermöglichen. Auch am 3'-Ende der transgen zu exprimierenden Nukleinsäuresequenzen können zusätzliche vorteilhafte Sequenzen insertiert werden, wie weitere regulatorische Elemente oder Terminatoren. Die transgen zu exprimierenden Nukleinsäuresequenzen können in einer oder mehreren Kopien im
- 45 Genkonstrukt enthalten sein.

Ferner ist es möglich, nach dem Startcodon eine sogenannte Downstream-Box einzufügen, die die Expression im Allgemeinen steigert (Translations-"Enhancer" WO 00/07431; WO 01/21782).

- 5 Als genetische Kontrollsequenzen v.a. bei der Kerntransformation geeignete Polyadenylierungssignale sind pflanzliche Polyadenylierungssignale, vorzugsweise solche, die im wesentlichen T-DNA Polyadenylierungssignale aus *Agrobacterium tumefaciens*, insbesondere des Gens 3 der T-DNA (Octopin Synthase) des Ti-Plasmids
- 10 pTiACHS entsprechen (Gielen et al., EMBO J. 3 (1984), 835 ff) oder funktionelle Äquivalente davon. Beispiele für besonders geeignete Terminatorsequenzen sind der OCS (Octopin-Synthase)-Terminator und der NOS (Nopalinsynthase)-Terminator.
- 15 Die erfindungsgemäßen Transformationsvektoren und Insertionssequenzen können weitere Nukleinsäuresequenzen umfassen. Solche Nukleinsäuresequenzen können bevorzugt Expressionskassetten darstellen. Beispielfhaft aber nicht einschränkend für die in den Expressionskonstrukten zu exprimierenden DNA-Sequenzen seien zu
- 20 nennen:

1. Selektionsmarker

- "Selektionsmarker" meint all solche Nukleinsäure- oder Proteinsequenzen, deren Expression (d.h. Transkription und ggf. Translation) einer Zelle, Gewebe oder Organismus einen anderen Phänotyp verleiht, als einer nicht transformierten. Selektionsmarker umfasst beispielsweise solche Nukleinsäure- oder Proteinsequenzen, deren Expression einer Zelle, Gewebe oder Organismus einen Vor-
- 30 teil (positiver Selektionsmarker) oder Nachteil (negativer Selektionsmarker) gegenüber Zellen vermittelt, die diese Nukleinsäure oder Protein nicht exprimieren. Positiven Selektionsmarkern wirken beispielsweise dadurch, dass eine auf die Zelle inhibitorisch wirkende Substanz detoxifiziert wird (Bsp. Antibiotika-/
- 35 Herbizidresistenz), oder eine Substanz gebildet wird, welche der Pflanze unter den gewählten Bedingungen verbesserte Regeneration oder erhöhtes Wachstum ermöglicht (zum Beispiel nutritive Marker, hormonproduzierender Marker wie ipt; s.u.). Eine andere Form positiver Selektionsmarker umfasst mutierte Proteine oder RNAs,
- 40 die gegenüber einem selektiven Agens nicht empfindlich sind (beispielsweise 16S rRNA Mutanten, die unempfindlich gegenüber Spectinomycin sind). Negative Selektionsmarker wirken beispielsweise dadurch, dass sie die Bildung einer toxischen Substanz in den transformierten Zellen katalysieren (zum Beispiel das
- 45 codA Gen). Ferner kann Selektionsmarker auch Reporterproteine umfassen, insofern diese geeignet sind, transformierte von nicht-

61

transformierten Zellen, Geweben oder Organen (beispielsweise durch Farbgebung oder einen anderen detektierbaren Phänotyp) zu unterscheiden.

- 5 Nachfolgende Selektionsmarker seien beispielhaft und nicht einschränkend zu nennen:

1.1 Positive Selektionsmarker:

- 10 Der mit der Expressionskassette in den Zellkern oder die Plastiden eingebrachte selektionierbare Marker verleiht den erfolgreich transformierten Zellen eine Resistenz gegen ein Biozid (zum Beispiel ein Herbizid wie Phosphinothricin, Glyphosat oder Bromoxynil), einen Metabolismusinhibitor wie 2-Desoxyglucose-6-phosphat
15 (WO 98/45456) oder ein Antibiotikum, wie zum Beispiel Tetracycline, Ampicillin, Kanamycin, G 418, Neomycin, Bleomycin oder Hygromycin. Der Selektionsmarker erlaubt die Selektion der transformierten Zellen von untransformierten (McCormick et al., Plant Cell Reports 5 (1986), 81-84; Dix PJ & Kavanagh TA (1995) Euphytica 85: 29-34).

Besonders bevorzugte Selektionsmarker sind solche, die eine Resistenz gegen Herbizide verleihen. Beispielhaft als Selektionsmarker seien genannt:

25

- DNA Sequenzen, die für Phosphinothricinacetyltransferasen (PAT) kodieren, welche die freie Aminogruppe des Glutaminsynthaseinhibitors Phosphinothricin (PPT) acetylieren und damit eine Detoxifizierung des PPT erreichen (de Block et al.
30 1987, EMBO J. 6, 2513-2518) (auch Bialaphos® resistenzgen (bar) genannt). Das bar Gen kodierend für eine Phosphinothricinacetyltransferase (PAT) kann aus beispielsweise Streptomyces hygroscopicus oder S. viridochromogenes isoliert werden. Entsprechende Sequenzen sind dem Fachmann bekannt
35 (aus Streptomyces hygroscopicus GenBank Acc.-No.: X17220 und X05822, aus Streptomyces viridochromogenes GenBank Acc.-No.: M 22827 und X65195; US 5,489,520). Ferner sind synthetische Gene beispielsweise für die Expression in Plastiden beschrieben. Ein synthetisches Pat Gen ist
40 beschrieben in Becker et al. (1994) The Plant J. 5:299-307. Die Gene verleihen Resistenz gegen das Herbizid Bialaphos oder Glufosinat und sind vielbenutzte Marker in transgenen Pflanzen (Vickers JE et al. (1996). Plant Mol Biol Reporter 14:363-368; Thompson CJ et al. (1987) EMBO J 6:2519-2523).

45

62

- 5-Enolpyruvylshikimat-3-phosphatsynthasegene (EPSP Synthasegene), die eine Resistenz gegen Glyphosat (N-(phosphonomethyl)glycin) verleihen. Das unselektive Herbizid Glyphosat hat die 5-Enolpyruvyl-3-phosphoshikimatsynthase (EPSPS) als molekulares Target. Diese hat eine Schlüsselfunktion in der Biosynthese aromatischer Aminosäuren in Mikroben und Pflanzen, jedoch nicht in Säugern (Steinrucken HC et al. (1980) Biochem. Biophys. Res. Commun. 94:1207-1212; Levin JG und Sprinson DB (1964) J. Biol. Chem. 239: 1142-1150; Cole DJ (1985) Mode of action of glyphosate a literature analysis, p. 48-74. In: Grossbard E und Atkinson D (eds.). The herbicide glyphosate. Butterworths, Boston.). Glyphosat-tolerante EPSPS Varianten werden bevorzugt als Selektionsmarker verwendet (Padgett SR et al. (1996). New weed control opportunities: development of soybeans with a Roundup Ready™ gene. In: Herbicide Resistant Crops (Duke, S.O., ed.), pp. 53-84. CRC Press, Boca Raton, FL; Saroha MK und Malik VS (1998) J Plant Biochemistry and Biotechnology 7:65-72). Das EPSPS Gen des Agrobakterium sp. strain CP4 hat eine natürliche Toleranz gegen Glyphosat, die auf entsprechende transgene Pflanzen transferiert werden kann. Das CP4 EPSPS Gen wurde aus Agrobakterium sp. strain CP4 kloniert (Padgett SR et al. (1995) Crop Science 35(5):1451-1461). Sequenzen von 5-Enolpyruvylshikimate-3-phosphate-synthasen, die Glyphosat-tolerant sind, wie beispielsweise beschrieben in US 5,510,471; US 5,776,760; US 5,864,425; US 5,633,435; US 5,627,061; US 5,463,175; EP 0 218 571, sind sowohl in den Patenten beschrieben als auch in der GenBank hinterlegt. Weitere Sequenzen sind beschrieben unter GenBank Accession X63374. Ferner ist das aroA Gen bevorzugt (M10947 S. typhimurium aroA locus 5-enolpyruvylshikimate-3-phosphate synthase (aroA protein) gene).
- das für das Glyphosat® degradierende Enzyme kodierende gox Gen (Glyphosatoxidoreduktase). GOX (beispielsweise die Glyphosatoxidoreduktase aus Achromobacter sp.) katalysiert die Spaltung einer C-N Bindung im Glyphosat, welches so zu Aminomethylphosphonsäure (AMPA) und Glyoxylat umgesetzt wird. GOX kann dadurch eine Resistenz gegen Glyphosat vermitteln (Padgett SR et al. (1996) J Nutr. 1996 Mar;126(3):702-16; Shah D et al. (1986) Science 233: 478-481).
- das deh Gen (kodierend für eine Dehalogenase, die Dalapon® inaktiviert), (GenBank Acc.-No.: AX022822, AX022820 sowie WO99/27116)

- 5 - bxn Gene, die für Bromoxynil degradierende Nitrilase-enzyme kodieren. Beispielsweise die Nitrilase aus *Klebsiella ozananae*. Sequenzen sind in der Genbank beispielsweise unter den Acc.-No: E01313 (DNA encoding bromoxynil specific nitrilase) und J03196 (*K. pneumoniae* bromoxynil-specific nitrilase (bxn) gene, complete cds) zu finden.
- 10 - Neomycinphosphotransferasen verleihen eine Resistenz gegen Antibiotika (Aminoglykoside) wie Neomycin, G418, Hygromycin, Paromomycin oder Kanamycin, indem sie durch eine Phosphorylierungsreaktion deren inhibierende Wirkung reduzieren. Besonders bevorzugt ist das nptII Gen. Sequenzen können aus der GenBank erhalten werden (AF080390 Mini-transposon mTn5-GNm; AF080389 Minitransposon mTn5-Nm, complete sequence). Zudem ist das Gen bereits Bestandteil zahlreicher Expressionsvektoren und kann unter Verwendung von dem Fachmann geläufigen Verfahren (wie beispielsweise Polymerasekettenreaktion) aus diesen isoliert werden (AF234316 pCAMBIA-2301; AF234315 pCAMBIA-2300, AF234314 pCAMBIA-2201). Das NPTII Gen kodiert für eine Aminoglycosid-3'-O-phosphotransferase aus *E.coli*, Tn5 (GenBank Acc.-No: U00004 Position 1401-2300; Beck et al. (1982) Gene 19 327-336). Darüber hinaus kann auch das aphA-6 Gen aus *Acinetobacter baumannii*, kodierend für eine Aminoglycosidphosphotransferase, als Selektionsmarker genutzt werden (Huang et al. (2002) Mol Genet Genomics 268:19-27).
- 30 - das DOG^R1-Gen. Das Gen DOG^R1 wurde aus der Hefe *Saccharomyces cerevisiae* isoliert (EP 0 807 836). Es codiert für eine 2-Desoxyglukose-6-phosphat Phosphatase, die Resistenz gegenüber 2-DOG verleiht (Randez-Gil et al. 1995, Yeast 11, 1233-1240; Sanz et al. (1994) Yeast 10:1195- 1202, Sequenz: GenBank Acc.-No.: NC001140 chromosom VIII, *Saccharomyces cervisiae* Position 194799-194056).
- 35 - Sulfonylurea- und Imidazolinon inaktivierende Acetolactat-synthasen, die eine Resistenz gegen Imidazolinon/Sulfonylurea-Herbizide verleihen. Beispielfhaft seien für Imidazolinon-Herbizide die Wirkstoffe Imazamethabenz-methyl, Imazamox, Imazapyr, Imazaquin, Imazethapyr zu nennen. Für Sulfonylharnstoff-Herbizide seien beispielsweise Amidosulfuron, Azimsulfuron, Chlorimuronethyl, Chlorsulfuron, Cinosulfuron, Imazosulfuron, Oxasulfuron, Prosulfuron, Rimsulfuron, Sulfosulfuron zu nennen. Dem Fachmann sind zahlreiche weitere Wirkstoffe der genannten Klassen bekannt. Geeignet sind Nukleinsäuresequenzen wie beispielsweise die unter der GenBank Acc.-No.: X51514 abgelegte Sequenz für das *Arabidopsis*

thaliana Csr 1.2 Gen (EC 4.1.3.18) (Sathasivan K et al. (1990) Nucleic Acids Res. 18(8):2188). Acetolactatesynthasen, die eine Resistenz gegen Imidazolinon-Herbizide verleihen, sind ferner beschrieben unter den GenBank Acc.-No.:

- 5
- a) AB049823 *Oryza sativa* ALS mRNA for acetolactate synthase, complete cds, herbicide resistant biotype
- b) AF094326 *Bassia scoparia* herbicide resistant acetolactate synthase precursor (ALS) gene, complete cds
- 10
- c) X07645 Tobacco acetolactate synthase gene, ALS SuRB (EC 4.1.3.18)
- d) X07644 Tobacco acetolactate synthase gene, ALS SuRA (EC 4.1.3.18)
- 15
- e) A19547 Synthetic nucleotide mutant acetolactate synthase
- f) A19546 Synthetic nucleotide mutant acetolactate synthase
- 20
- g) A19545 Synthetic nucleotide mutant acetolactate synthase
- h) I05376 Sequence 5 from Patent EP 0257993
- 25
- i) I05373 Sequence 2 from Patent EP 0257993
- j) AL133315
- 30 - Hygromycinphosphotransferasen (X74325 *P. pseudomallei* gene for hygromycin phosphotransferase) die eine Resistenz gegen das Antibiotikum Hygromycin verleihen. Das Gen ist Bestandteil zahlreicher Expressionsvektoren und kann unter Verwendung von dem Fachmann geläufigen Verfahren (wie beispielsweise Polymerasekettenreaktion) aus diesen isoliert werden
- 35 (AF294981 pINDEX4; AF234301 pCAMBIA-1380; AF234300 pCAMBIA-1304; AF234299 pCAMBIA-1303; AF234298 pCAMBIA-1302; AF354046 pCAMBIA-1305.; AF354045 pCAMBIA-1305.1)
- 40 - Resistenzgene gegen
- a) Chloramphenicol (Chloramphenicolacetyltransferase),
- b) Tetracyclin, verschiedene Resistenzgene sind beschrieben
- 45 z.B. X65876 *S. ordonez* genes class D tetA and tetR for tetracycline resistance and repressor proteins
- X51366 *Bacillus cereus* plasmid pBC16 tetracycline

resistance gene. Zudem ist das Gen bereits Bestandteil zahlreicher Expressionsvektoren und kann unter Verwendung von dem Fachmann geläufigen Verfahren (wie beispielsweise Polymerasekettenreaktion) aus diesen isoliert werden

5

- c) Streptomycin, verschiedene Resistenzgene sind beschrieben z.B. mit der GenBank Acc.-No.:AJ278607 *Corynebacterium acetoacidophilum* ant gene for streptomycin adenylyl-transferase.

10

- d) Zeocin, das entsprechende Resistenzgen ist Bestandteil zahlreicher Klonierungsvektoren (z.B. L36849 Cloning vector pZEO) und kann unter Verwendung von dem Fachmann geläufigen Verfahren (wie beispielsweise Polymerasekettenreaktion) aus diesen isoliert werden.

15

- e) Ampicillin (β -Lactamase Gen; Datta N, Richmond MH. (1966) *Biochem J.* 98(1):204-9; Heffron F et al (1975) *J. Bacteriol* 122: 250-256; das Amp Gen wurde zuerst zur Herstellung des *E. coli* Vektors pBR322 kloniert; Bolivar F et al. (1977) *Gene* 2:95-114). Die Sequenz ist Bestandteil zahlreicher Klonierungsvektoren und kann unter Verwendung von dem Fachmann geläufigen Verfahren (wie beispielsweise Polymerasekettenreaktion) aus diesen isoliert werden.

20

25

- Gene wie die Isopentenyltransferase aus *Agrobacterium tumefaciens* (strain:PO22) (Genbank Acc.-No.: AB025109). Das *ipt* Gen ist ein Schlüsselenzym der Cytokinin-Biosynthese. Seine Überexpression erleichtert die Regeneration von Pflanzen (z.B. Selektion auf Cytokinin-freiem Medium). Das Verfahren zur Nutzung des *ipt* Gens ist beschrieben (Ebinuma H et al. (2000) *Proc Natl Acad Sci USA* 94:2117-2121; Ebinuma, H et al. (2000) Selection of Marker-free transgenic plants using the oncogenes (*ipt*, *rol A*, *B*, *C*) of *Agrobacterium* as selectable markers, In *Molecular Biology of Woody Plants*. Kluwer Academic Publishers).

30

35

Verschiedene weitere positive Selektionsmarker, die den transformierten Pflanzen einen Wachstumsvorteil gegenüber nicht-transformierten verleihen, sowie Verfahren zu ihrer Verwendung sind u.a. beschrieben in EP-A 0 601 092. Beispielhaft sind zu nennen β -Glucuronidase (in Verbindung mit z.B. Cytokininglucuronid), Mannose-6-phosphat-Isomerase (in Verbindung mit Mannose), UDP-Galaktose-4-Epimerase (in Verbindung mit z.B. Galactose), wobei

40

45

Mannose-6-phosphat-Isomerase in Verbindung mit Mannose besonders bevorzugt ist.

- Für einen in Plastiden funktionellen Selektionsmarker sind insbesondere solche bevorzugt, die eine Resistenz gegen Spectinomycin, Streptomycin, Kanamycin, Lincomycin, Gentamycin, Hygromycin, Methotrexat, Bleomycin, Phleomycin, Blastidicin, Sulfonamid,, Phosphinotricin, Chlorsulfuron, Bromoxymil, Glyphosat, 2,4-Datrazin, 4-methyltryptophan, Nitrat, S-aminoethyl-L-cysteine, Lysin/Threonin, Aminoethyl-Cystein oder Betainaldehyd verleihen. Besonders bevorzugt sind die Gene *aadA*, *nptII*, *BADH*, *FLARE-S* (eine Fusion aus *aadA* und GFP, beschrieben bei Khan MS & Maliga P, 1999 Nature Biotech 17: 910-915).
- Als in Plastiden funktionelle Selektionsmarker ist vor allem das *aadA* Gen beschrieben (Svab Z und Maliga P (1993) Proc Natl Acad Sci USA 90:913-917). Ferner beschrieben sind modifizierte 16S rDNA, das *nptII* Gen (Kanamycinresistenz) und das *bar* Gen (Phosphinotricinresistenz). Aufgrund der Bevorzugung des Selektionsmarker *aadA* wird dieser bevorzugt "recycled" d.h. nach seiner Verwendung aus dem Genom bzw. Plastom deletiert (Fischer N et al. (1996) Mol Gen Genet 251:373-380; Corneille S et al. (2001) Plant J 27:171-178), so dass *aadA* in weiteren Transformationen einer bereits transplastomen Pflanze wieder als Selektionsmarker benutzt werden kann. Als weiterer möglicher Selektionsmarker ist die Betaine-aldehyd-Dehydrogenase (*BADH*) aus Spinat beschrieben (Daniell H et al. (2001) Trends Plant Science 6:237-239; Daniell H et al. (2001) Curr Genet 39:109-116; WO 01/64023; WO 01/64024; WO 01/64850). Auch lethal wirkende Agenzen wie beispielsweise Glyphosat können in Verbindung mit entsprechend detoxifizierenden oder resistenten Enzymen genutzt werden (WO 01/81605).

- Es können auch *binding type marker* genutzt werden. Beispielsweise, können an einer Stelle der 23S rDNA (Position 2073 oder 2074 der 23S rRNA aus Tabak, Sequenz: AAAGACCCTATGAAG) Punktmutationen eingeführt werden (z.B. Sequenz: GGAGACCCTATGAAG), die den aus einer so mutierten 23S rDNA abgeleiteten Ribosomen Resistenz gegenüber Lincomycin vermitteln (Cseplö A et al. (1988) Mol Gen Genet 214:295-299). Weitere Punktmutation umfassen solche in der 16S rRNA von Tabak, die Resistenz gegenüber Spectinomycin verleihen (Mutation unterstrichen):

a) 5'-GGAAGGTGAGGATGC-3' ("nativ" hier: A)

- Andere Mutationen in die 16S rRNA verleihen eine Resistenz gegen Streptomycin:

b) 5'-GAATGAAACTA-3' ("nativ" hier: C)

1.2 Negative Selektionsmarker

- 5 Negative Selektionsmarker ermöglichen beispielsweise die Selektion von Organismen mit erfolgreich deletierten Sequenzen, die das Markergen umfassen (Koprek T et al. (1999) The Plant Journal 19(6):719-726). Beispielsweise können für Rekombinasen, Selektionsmarker oder für DSBI-Enzyme kodierende Sequenzen nach
- 10 erfolgreicher Anwendung des erfindungsgemäßen Verfahrens wieder aus dem Genom/Plastom deletiert werden.

- Bei der negativen Selektion wird beispielsweise durch den in die Pflanze eingebrachten negativen Selektionsmarker eine Verbindung,
- 15 die ansonsten für die Pflanze keine nachteilige Wirkung hat, in eine Verbindung mit nachteiliger Wirkung umgesetzt. Ferner sind Gene geeignet, die per se eine nachteilige Wirkung haben, wie zum Beispiel TK thymidine kinase (TK) und Diphtheria Toxin A Fragment (DT-A), das codA Genprodukt kodierend für eine Cytosindeaminase
- 20 (Gleave AP et al. (1999) Plant Mol Biol 40(2):223-35; Perera RJ et al. (1993) Plant Mol Biol 23(4): 793-799; Stougaard J (1993) Plant J 3:755-761), das Cytochrom P450 Gen (Koprek et al. (1999) Plant J. 16:719-726), Gene kodierend für eine Haloalkan Dehalogenase (Naested H (1999) Plant J. 18:571-576), das iaaH Gen
- 25 (Sundaresan V et al. (1995) Genes & Development 9:1797-1810) oder das tms2 Gen (Fedoroff NV & Smith DL (1993) Plant J 3:273-289).

- Die jeweils für die Selektion verwendeten Konzentrationen der Antibiotika, Herbizide, Biozide oder Toxine müssen an die
- 30 jeweiligen Testbedingungen bzw. Organismen angepasst werden. Beispielfhaft seien für Pflanzen zu nennen Kanamycin (Km) 50 mg/L, Hygromycin B 40 mg/L, Phosphinothricin (Ppt) 6 mg/L, Sepctinomycin (Spec) 500 mg/L.

- 35 Ferner können funktionelle Analoga der genannten Nukleinsäuren kodierend für Selektionsmarker exprimiert werden. Funktionelle Analoga meint hier all die Sequenzen, die im wesentlichen die gleiche Funktion haben d.h. zu einer Selektion transformierter Organismen befähigt sind. Dabei kann das funktionelle Analogon
- 40 sich in anderen Merkmalen durchaus unterscheiden. Es kann zum Beispiel eine höhere oder niedrigere Aktivität haben oder auch über weitere Funktionalitäten verfügen.

- Funktionelle Analoga meint ferner Sequenzen, die für Fusions-
- 45 proteine kodieren bestehend aus einem der bevorzugten Selektionsmarker und einem anderen Protein zum Beispiel einem weiteren bevorzugten Selektionsmarker, einem der unten genannten Reporter-

proteine oder einer PLS. Beispielhaft sei eine Fusion des GFP (grün fluoreszierenden Proteins) und des aadA Gens genannt (Sidorov VA et al. (1999) Plant J 19:209-216).

5 2. Reportergene

- Reportergene kodieren für leicht quantifizierbare Proteine, die über Eigenfarbe oder Enzymaktivität eine Bewertung der Transformationseffizienz, des Expressionsortes oder -zeitpunktes oder
- 10 die Identifizierung transgener Pflanzen gewährleisten. Ganz besonders bevorzugt sind dabei Gene kodierend für Reporter-Proteine (siehe auch Schenborn E, Groskreutz D. Mol Biotechnol. 1999; 13(1):29-44) wie
- 15 - "green fluorescence protein" (GFP) (Chui WL et al., Curr Biol 1996, 6:325-330; Leffell SM et al., Biotechniques. 23(5):912-8, 1997; Sheen et al.(1995) Plant Journal 8(5):777-784; Haseloff et al.(1997) Proc Natl Acad Sci USA 94(6):2122-2127; Reichel et al.(1996) Proc Natl Acad Sci USA 93(12):5888-5893; Tian et al. (1997) Plant Cell Rep 16:267-271; WO 97/41228).
- 20 - Chloramphenicoltransferase,
- 25 - Luziferase (Millar et al., Plant Mol Biol Rep 1992 10:324-414; Ow et al. (1986) Science, 234:856-859); erlaubt Biolumineszenzdetektion.
- β -Galactosidase, kodiert für ein Enzym für das verschiedenen
- 30 chromogene Substrate zur Verfügung stehen.
- β -Glucuronidase (GUS) (Jefferson et al., EMBO J. 1987, 6, 3901-3907) oder das uidA Gen, das ein Enzym für verschiedene chromogene Substrate kodiert.
- 35 - R-Locus Genprodukt:Protein, das die Produktion von Anthocyaninpigmenten (rote Färbung) in pflanzlichen Gewebe reguliert und so eine direkte Analyse der Promotoraktivität ohne Zugabe zusätzlicher Hilfsstoffe oder chromogener Substrate
- 40 ermöglicht (Dellaporta et al., In: Chromosome Structure and Function: Impact of New Concepts, 18th Stadler Genetics Symposium, 11:263-282, 1988).
- β -Lactamase (Sutcliffe (1978) Proc Natl Acad Sci USA 75:3737-3741), Enzym für verschiedene chromogene Substrate
- 45 (z.B. PADAC, eine chromogenes Cephalosporin).

69

- xylE Genprodukt (Zukowsky et al. (1983) Proc Natl Acad Sci USA 80:1101-1105), Catecholdioxygenase, die chromogene Catechole umsetzen kann.
- 5 - Alpha-Amylase (Ikuta et al. (1990) Bio/technol. 8:241-242).
- Tyrosinase (Katz et al. (1983) J Gen Microbiol 129:2703-2714), Enzym, das Tyrosin zu DOPA und Dopaquinon oxidiert, die infolge das leicht nachweisbare Melanin bilden.
- 10 - Aequorin (Prasher et al. (1985) Biochem Biophys Res Commun 126(3):1259-1268), kann in der Calcium-sensitiven Biolumineszenzdetektion verwendet werden.
- 15 Der Selektionsmarker bzw. das Reportergen ist bevorzugt auf dem RE- bzw. RE/DSB-Konstrukt und/oder Transformationskonstrukt, besonders bevorzugt auf der Insertionssequenz kodiert. Er kann aber auch auf einem unabhängigen Transformationskonstrukt kodiert sein, welches in einer Co-Transformation mit dem Transformations-
- 20 konstrukt von Interesse in den Kern oder die Plastiden einer Pflanzenzelle eingebracht wird.

Die erfindungsgemäßen Transformationsvektoren und Insertionssequenzen können weiteren Funktionselemente enthalten. Der

- 25 Begriff der weiteren Funktionselemente ist breit zu verstehen. Bevorzugt sind all solche Elemente gemeint, die einen Einfluss auf Herstellung, Vermehrung, Funktion, Nutzen oder Wert der im Rahmen des erfindungsgemäßen Verfahrens zum Einsatz kommenden Insertionssequenzen, Transformationskonstrukte- oder -vektoren
- 30 haben. Beispielhaft jedoch nicht einschränkend seien für die weiteren Funktionselemente zu nennen:
 - i. Replikationsursprünge (ORI; origin of DNA replication), die eine Vermehrung der erfindungsgemäßen Expressionskassetten
 - 35 oder Vektoren in zum Beispiel E. coli oder aber auch in Plastiden gewährleisten. Beispielhaft für E.coli ORIs seien genannt der pBR322 ori, der P15A ori (Sambrook et al.: Molecular Cloning. A Laboratory Manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989) oder der
 - 40 colE1-ORI beispielsweise aus pBLUESCRIPT. Plastidäre ORI sind beschrieben in US 5,693,507, US 5,932,479 oder WO 99/10513.
 - ii. Multiple Klonierungsregionen (MCS) erlauben und erleichtern die Insertion eines oder mehrerer Nukleinsäuresequenzen.

- iii. Sequenzen, die eine homologe Rekombination bzw. Insertion in das Genom bzw. Plastom eines Wirtsorganismus ermöglichen.
- iv. Elemente zum Beispiel "Bordersequenzen", die einen Agrobakterien-vermittelte Transfer in Pflanzenzellen für die Übertragung und Integration ins Pflanzengenom ermöglichen, wie zum Beispiel die rechte oder linke Begrenzung der T-DNA oder die vir-Region.
- 10 Die Einführung einer Insertionssequenz oder eines Expressionskonstruktes - beispielsweise für eine Rekombinase, ein DSBI-Enzym oder einen Selektionsmarker - kann vorteilhaft unter Verwendung von Vektoren realisiert werden, in die diese Konstrukte bzw. Kassetten insertiert werden. Vektoren können beispielhaft
- 15 Plasmide, Cosmide, Phagen, Viren, Retroviren oder auch Agrobakterien sein.

- In einer vorteilhaften Ausführungsform wird die Einführung der Expressionskassette mittels Plasmidvektoren realisiert. Bevorzugt sind solche Vektoren, die eine stabile Integration der Expressionskassette in das Wirtsgenom bzw. -plastom ermöglichen.
- 20

- Die Herstellung eines transformierten Organismus oder einer transformierten Zelle erfordert, dass die entsprechende DNA
- 25 in die entsprechende Wirtszelle bzw. in die Plastiden derselben eingebracht wird. Für diesen Vorgang, der als Transformation bezeichnet wird, steht eine Vielzahl von Methoden zur Verfügung (siehe auch Keown et al. (1990) Methods in Enzymology 185:527-537). So kann die DNA beispielhaft direkt durch Mikroinjektion, Elektroporation oder durch Bombardierung mit DNA-be-
- 30 schichteten Mikropartikeln eingeführt werden. Auch kann die Zelle chemisch, zum Beispiel mit Polyethylenglycol, permeabilisiert werden, so dass die DNA durch Diffusion in die Zelle gelangen kann. Die Transformation kann auch durch Fusion mit anderen DNA-
- 35 enthaltenden Einheiten wie Minicells, Zellen, Lysosomen oder Liposomen erfolgen. Zu nennen sind ferner die Transfektion unter Einsatz von Calciumphosphat, DEAE-Dextran oder kationischen Lipiden, Transduktion, Infektion, die Inkubation trockener Embryonen in DNA-haltiger Lösung, die Sonikation sowie die Trans-
- 40 formation intakter Zellen oder Gewebe durch Mikro- oder Makroinjektion in Gewebe oder Embryonen, Gewebeelektroporation oder die Vakuuminfiltration von Samen. Derartige Verfahren sind dem Fachmann geläufig. Im Falle von Injektion oder Elektroporation von DNA in pflanzliche Zellen sind keine besonderen Anforderungen
- 45 an das verwendete Plasmid gestellt. Einfache Plasmide wie die der pUC-Reihe können verwendet werden. Sollen vollständige Pflanzen aus den transformierten Zellen regeneriert werden, so ist es

nützlich, das sich auf dem Plasmid ein zusätzliches selektionierbares Markergen befindet. Auch die Verfahren zur Regeneration von Pflanzen aus Pflanzengeweben oder Pflanzenzellen sind beschrieben.

- 5 Es gibt verschiedene Möglichkeiten, die DNA in die Plastiden einzuschleusen. Für die vorliegende Erfindung ist lediglich entscheidend, dass die DNA in die Plastiden eingebracht wird. Die vorliegende Erfindung ist jedoch nicht limitiert auf ein
- 10 bestimmtes Verfahren. Jedes Verfahren, das ein Einbringen der zu transformierenden DNA in die Plastiden einer höheren Pflanze erlaubt, ist geeignet. Die stabile Transformation von Plastiden ist ein dem Fachmann geläufiges Verfahren und wurde für höhere Pflanzen beschrieben (u.a. bei Svab Z und Maliga P (1993) Proc
- 15 Natl Acad Sci USA 90(3):913-917). Die Methoden beruhen zum Beispiel auf einer Transformation mittels "Particle Gun" und einer Insertion in das plastidäre Genom durch homologe Rekombination unter Selektionsdruck. Weitere Methoden sind beschrieben in US 5,877,402. In EP-A 0 251 654 wird die DNA mittels Agro-
- 20 bakterium tumefaciens eingeführt (s. De Block M et al. (1985) EMBO J 4:1367-1372; Venkateswarlu K und Nazar RN (1991) Bio/Technology 9:1103-1105). Ferner wurde gezeigt, dass man mittels Elektroporation DNA in isolierte Chloroplasten einbringen und so eine transiente Expression erzielen kann (To KY et al. (1996)
- 25 Plant J 10:737-743). Bevorzugt ist eine Transformation mittels eines direkten Transfers der DNA in Plastiden von Protoplasten beispielsweise unter Verwendung von PEG (Polyethylenglykol) (Koop HU et al. (1996) Planta 199:193-201; Kofer W et al. (1998) In Vitro Cell Dev Biol Plant 34:303-309; Dix PJ und Kavanagh TA (1995)
- 30 Euphytica. 85:29-34; EP-A 0 223 247). Am meisten bevorzugt sind biolistische Transformationsverfahren. Dabei wird die zu transformierende DNA auf z.B. Gold- oder Wolframpartikel aufgebracht. Diese Partikel werden anschließend auf das zu transformierende Explantat beschleunigt (Dix PJ und Kavanagh TA (1995) Euphytica.
- 35 85:29-34; EP-A 0 223 247). Anschließend werden in der dem Fachmann geläufigen Weise transplastome Pflanzen unter Selektionsdruck auf geeignetem Medium regeneriert. Entsprechende Verfahren sind beschrieben (z.B. US 5,451,513; US 5,877,402; Svab Z et al. (1990) Proc Natl Acad Sci USA 87:8526-8530; Svab Z und Maliga P
- 40 (1993) Proc Natl Acad Sci USA 90:913-917). Darüber hinaus kann die DNA mittels Mikroinjektion in die Plastiden eingebracht werden. Ein besonderes Verfahren der Mikroinjektion wurde kürzlich beschrieben (Knoblauch M et al. (1999) Nature Biotech 17:906-909; van Bel AJE et al. (2001) Curr Opin Biotechnol
- 45 12:144-149). Dieses Verfahren ist für die vorliegende Erfindung besonders bevorzugt. Es ist auch möglich, durch Protoplastenfusion die Plastiden aus einer Art in eine andere Art einzu-

bringen, diese dort zu transformieren und anschließend durch Protoplastenfusion diese wieder in die ursprüngliche Art zu überführen (WO 01/70939).

- 5 Neben diesen "direkten" Transformationstechniken kann eine Transformation auch durch bakterielle Infektion mittels *Agrobacterium tumefaciens* oder *Agrobacterium rhizogenes* durchgeführt werden (Horsch RB (1986) Proc Natl Acad Sci USA 83(8):2571-2575; Fraley et al. (1983) Proc Natl Acad Sci USA 80:4803-4807; Bevan et al. (1983) Nature 304:184-187). Die Expressionskassette beispielsweise für das DSBII-Enzym wird bevorzugt in spezielle Plasmide integriert, entweder in einen Zwischenvektor (englisch: shuttle or intermediate vector) oder einen binären Vektor. Bevorzugt werden binäre Vektoren verwendet. Binäre Vektoren können sowohl
- 15 in *E. coli* als auch in *Agrobacterium* replizieren und können direkt in *Agrobacterium* transformiert werden (Holsters et al. (1978) Mol Gen Genet 163:181-187). Verschiedene binäre Vektoren sind bekannt und teilweise kommerziell erhältlich wie zum Beispiel pBIN19 (Clontech Laboratories, Inc. USA; Bevan et al. (1984) Nucl Acids Res 12:8711). Das Selektionsmarkergen erlaubt eine
- 20 Selektion transformierter *Agrobacteria* und ist zum Beispiel das nptII Gen, das eine Resistenz gegen Kanamycin verleiht. Das Binärplasmid kann beispielsweise durch Elektroporation oder andere Transformationsmethoden in den *Agrobacterienstamm*
- 25 übertragen werden (Mozo & Hooykaas 1991, Plant Mol. Biol. 16, 917-918). Die Kokultur der pflanzlichen Explantate mit dem *Agrobacterienstamm* findet in der Regel für zwei bis drei Tage statt. Das in diesem Fall als Wirtsorganismus fungierende *Agrobacterium* sollte bereits ein Plasmid mit der vir-Region ent-
- 30 halten. Viele Stämme von *Agrobacterium tumefaciens* sind in der Lage, genetisches Material zu übertragen, wie z.B. die Stämme EHA101[pEHA101] (Hood EE et al. (1996) J Bacteriol 168(3):1291-1301), EHA105[pEHA105] (Hood et al. (1993) Transgenic Research 2:208-218), LBA4404[pAL4404] (Hoekema et al. (1983) Nature 303:179-181), C58C1[pMP90] (Koncz and Schell (1986) Mol Gen Genet 204:383-396) und C58C1[pGV2260] (Deblaere et al. (1985) Nucl Acids Res 13: 4777-4788).

- Für den Transfer der DNA auf die pflanzliche Zelle werden pflanz-
- 40 liche Explantate mit *Agrobacterium tumefaciens* oder *Agrobacterium rhizogenes* kokultiviert. Ausgehend von infiziertem Pflanzenmaterial (z.B. Blatt-, Wurzel- oder Stengelteile, aber auch Protoplasten oder Suspensionen von Pflanzenzellen) können ganze Pflanzen unter Verwendung eines geeigneten Mediums, dass zum
 - 45 Beispiel Antibiotika oder Biozide zur Selektion transformierten Zellen enthalten kann, regeneriert werden. Ein mittransformierter Selektionsmarker erlaubt die Selektion von transformierten Zellen

- von untransformierten (McCormick et al. (1986) Plant Cell Reports 5:81-84). Die erhaltenen Pflanzen können in üblicher Weise gezüchtet, geselbstet und gekreuzt werden. Zwei oder mehr Generationen sollten kultiviert werden, um sicherzustellen, dass die genomische Integration stabil und vererblich ist. Die genannten Verfahren sind beispielsweise beschrieben bei Jenes B et al. (1993) Techniques for Gene Transfer, in: Transgenic Plants, Vol. 1, Engineering and Utilization, herausgegeben von Kung SD und Wu R, Academic Press, S. 128-143 sowie bei Potrykus (1991) Ann Rev Plant Physiol Plant Molec Biol 42:205-225).

Die Agrobakterium-vermittelte Transformation ist am besten für dicotyledone Pflanzenzellen geeignet, wohingegen die direkten Transformationstechniken sich für jeden Zelltyp eignen.

- 15 Die Agrobakterium-vermittelte Transformation wird besonders bevorzugt für die Kerntransformation, die direkten Transformationstechniken besonders bevorzugt für die Plastidentransformation eingesetzt.
- 20 Sobald eine überwiegend homotransplastome Pflanzenzelle nach dem erfindungsgemäßen Verfahren hergestellt wurde, kann eine vollständige Pflanze unter Verwendung von dem Fachmann bekannten Verfahren erhalten werden. Hierbei geht man beispielhaft von Kalluskulturen aus. Aus diesen noch undifferenzierten Zellmassen kann die Bildung von Spross und Wurzel in bekannter Weise induziert werden. Die erhaltenen Sprösslinge können ausgepflanzt und gezüchtet werden.

30 Deletionsverfahren

- Bei den beschriebenen erfindungsgemäßen Verfahren ist es auf verschiedenen Stufen vorteilhaft, bestimmte zuvor eingeführte Sequenzen (beispielsweise für Selektionsmarker, Rekombinasen und/oder DSBI-Enzyme) wieder aus dem Plastom oder Genom der Pflanze oder Zelle zu entfernen. So ist es von Vorteil aber nicht zwingend erforderlich, den Selektionsmarker, der beispielsweise bei der Insertion einer nicht-natürlichen RE-Sequenz und/oder DBS-Erkennungssequenz eingeführt wurde, aus der Masterpflanze wieder zu entfernen, weil dann in einer nachfolgenden Transformation (z.B. mit der Insertionssequenz) wieder der gleiche Selektionsmarker genutzt werden kann. Eine Deletion ist insbesondere vorteilhaft, da der Selektionsmarker nach der Selektionsphase nicht mehr unbedingt erforderlich und daher überflüssig ist. Die Deletion erhöht zudem die Verbraucherakzeptanz und ist unter zulassungstechnischen Überlegungen wünschenswert. Zudem wird der Proteinsyntheseapparat des Plastides nicht unnötig

durch die Synthese des Markerproteins belastet, was sich potentiell vorteilhaft auf die Eigenschaften der entsprechenden Pflanze auswirkt.

- 5 Zur gezielten Deletion von Sequenzen sind dem Fachmann verschiedenen Verfahren bekannt. Zu nennen ist beispielsweise jedoch nicht einschränkend die Excision mittels Rekombinasen. Verschiedene sequenzspezifische Rekombinationssystemen sind beschrieben wie das Cre/lox-System des Bacteriophagen P1 (Dale EC
10 und Ow DW (1991) Proc Natl Acad Sci USA 88:10558-10562; Russell SH et al. (1992) Mol Gen Genet 234: 49-59; Osborne BI et al. (1995) Plant J. 7, 687-701), das FLP/FRT System der Hefe (Kilby NJ et al. (1995) Plant J 8:637-652; Lyznik LA et al. (1996) Nucleic Acids Res 24:3784-3789), die Gin Rekombinase des Mu Phagen,
15 die Pin Rekombinase aus E. coli oder das R/RS System des pSR1 Plasmids (Onouchi H et al. (1995) Mol Gen Genet 247:653-660; Sugita K et al. (2000) Plant J 22:461-469). Diese Verfahren sind nicht nur für die Deletion von DNA Sequenzen aus dem Kerngenom, sondern auch aus dem Plastom nutzbar (Corneille et al. (2001)
20 Plant J 27: 171-178; Hajdukiewicz et al. (2001) Plant J 27:161-170). Als weitere Rekombinasen können beispielsweise eingesetzt werden: PhiC31 (Kuhstoss & Rao (1991) J Mol Biol 222:897-908), TP901-1 (Christiansen et al. (1996) J Bacteriol 178:5164-5173), xisF aus Anabaena (Ramaswamy et al. (1997) Mol
25 Microbiol 23:1241-1249), Integrase vom Phagen PhiLC3 (Lillehaug et al. (1997) Gene 188:129-136) oder die Rekombinase kodiert vom sre Gen des R4-Phagen (Matsuura et al. (1996) J Bacteriol 178:3374-3376).
- 30 In einer bevorzugten Ausführungsform wird die Deletion jedoch durch intramolekulare (z.B. intrachromosomale) Rekombination aufgrund von entsprechend eingebrachten Sequenzduplikationen realisiert. Letzteres kann durch das gezielte Einführen von Doppelstrangbrüchen nahe der Sequenz-Duplikationen in seiner Effizienz
35 gesteigert werden (vgl. Fig. 8). Dazu wird die zu deletierende Sequenz beidseitig von Homologiesequenzen H1 und H2 flankiert, die eine ausreichende Länge und Homologie aufweisen, um miteinander zu rekombinieren. Die Rekombination wird durch die Induktion mindestens eines sequenzspezifischen Doppelstrangbruchs
40 nahe einer der beiden Homologiesequenzen gelegenen DSB-Erkennungssequenz induziert. Bevorzugt ist diese DSB-Erkennungssequenz zwischen den beiden Homologiesequenzen lokalisiert. Zur Induktion des Doppelstrangbruchs wird bevorzugt ein DSB-I-Enzym exprimiert oder eingebracht. Besonders bevorzugt wird dieses Verfahren zur
45 Deletion von Selektionsmarkern aus dem Plastom genutzt.

75

Ein anderer Gegenstand der Erfindung betrifft die mit dem erfindungsgemäßen Verfahren hergestellten transplastomen, überwiegend homoplastomen Pflanzen, sowie Teile derselben wie Blätter, Wurzeln, Samen, Früchte, Knollen, Pollen oder Zell-

5 kulturen, Kallus usw. - abgeleitet von solchen.

Ein anderer Gegenstand der Erfindung betrifft die in dem erfindungsgemäßen Verfahren zum Einsatz kommenden Pflanzen, die eine erfindungsgemäße Expressionskassette für ein Fusion-

10 sprotein aus einer PLS und einer Rekombinase der Familie der Resolvasen/Invertasen unter Kontrolle eines im pflanzlichen Zellkern funktionellen Promotors enthalten. Besonders bevorzugt liegt die Expressionskassette stabil integriert in die nukleäre DNA vor. Umfasst sind ferner Teile derselben wie Blätter, Wurzeln,

15 Samen, Knollen, Früchte, Pollen oder Zellkulturen, Kallus usw. - abgeleitet von solchen Pflanzen.

Ein weiterer Gegenstand der Erfindung betrifft Expressionskassetten, die Nukleinsäuresequenzen kodierend für Rekombinasen

20 der Familie der Resolvasen/Invertasen unter Kontrolle eines in pflanzlichen Plastiden funktionellen Promotors enthalten. Besonders bevorzugt liegt die Expressionskassette stabil integriert in das Plastom vor. Umfasst sind ferner Teile derselben wie Blätter, Wurzeln, Samen, Knollen, Früchte, Pollen oder Zell-

25 kulturen, Kallus usw. - abgeleitet von solchen Pflanzen.

Von Menschen und Tieren verzehrbare erfindungsgemäße, genetisch veränderte Pflanzen können auch beispielsweise direkt oder nach an sich bekannter Aufbereitung als Nahrungsmittel oder Futter-

30 mittel verwendet werden.

Ein weiterer Gegenstand der Erfindung betrifft die Verwendung der oben beschriebenen erfindungsgemäßen transplastomen, überwiegend homoplastomen Pflanzen und der von ihnen abgeleiteten Zellen,

35 Zellkulturen, Teile - wie zum Beispiel bei transgenen pflanzlichen Organismen Wurzeln, Blätter etc.-, und transgenes Vermehrungsgut wie Samen oder Früchte, zur Herstellung von Nahrungs- oder Futtermitteln, Pharmazeutika oder Feinchemikalien.

40 Feinchemikalien meint Enzyme wie beispielsweise die unten genannten industriellen Enzyme, Vitamine wie beispielsweise Tocopherolen und Tocotrienolen (z.B. Vitamin E) und Vitamin B2, Aminosäuren wie beispielsweise Methionin, Lysin oder Glutamat, Kohlenhydrate wie beispielsweise Stärke, Amylose, Amylopektin

45 oder Saccharose, Fettsäuren wie beispielsweise gesättigte, ungesättigte und polyungesättigte Fettsäuren, natürliche und synthetische Geschmacks-, Aromastoffe wie beispielsweise Linalo-

76

ol, Menthol, Borneon (Kampfer), Pinen, Limonen oder Geraniol und Farbstoffe wie beispielsweise Retinoide (z.B. Vitamin A), Flavonoide (z.B. Quercetin, Rutin, Tangeretin, Nobiletin) oder Carotinoide (z.B. β -Carotin, Lycopin, Astaxanthin). Besonders
5 bevorzugt ist die Produktion von Tocopherolen und Tocotrienolen sowie Carotinoiden. Die Züchtung der transformierten Wirtsorganismen sowie die Isolierung aus den Wirtsorganismen bzw. aus dem Züchtungsmedium erfolgt mit dem Fachmann bekannten Verfahren. Die Produktion von Pharmazeutika, wie zum Beispiel Antikörpern
10 oder Vakzzinen ist beschrieben (Hood EE, Jilka JM. (1999) Curr Opin Biotechnol. 10(4):382-386; Ma JK und Vine ND (1999) Curr Top Microbiol Immunol. 236:275-92).

Insbesondere eignet sich das erfindungsgemäße Verfahren zur Herstellung von industriellen Enzymen im Rahmen eines sogenannten
15 "Phytofarming". Beispielhaft jedoch nicht einschränkend seien für die industriellen Enzyme zu nennen Lipasen, Esterasen, Proteasen, Nitrilasen, Acylasen, Epoxyhydrolasen, Amidasen, Phosphatasen, Xylanasen, Alkoholdehydrogenasen, Amylasen, Glucosidasen,
20 Galactosidasen, Pullulanasen, Endocellulasen, Glucanasen, Cellulasen, Nukleasen, Chitindeacetylase, Monoaminoxidasen, Lysozyme und Laccasen.

Im Rahmen der Erfindung insbesondere bevorzugte Ausführungsformen
25 werden unten im Rahmen der Erläuterungen zu den Abbildungen näher beschrieben.

Sequenzen

- 30 1. SEQ ID NO: 1
pCB42-94 Basisvektor für die Plastidentransformation.
2. SEQ ID NO:2
In die Multiple Klonierungsstelle von pCB42-94 (SEQ ID NO: 1)
35 insertierte Nukleinsäuresequenz. Resultierender Vektor:
pCB199-3.
3. SEQ ID NO:3
In die Multiple Klonierungsstelle von pCB42-94 (SEQ ID NO: 1)
40 insertierte Nukleinsäuresequenz. Resultierender Vektor:
pCB401-20
4. SEQ ID NO: 4
Synthetische Sequenz der Homing-Endonuklease I-PpoI (ORF: 16
45 bis 507)

5. SEQ ID NO: 5
Proteinsequenz der Homing-Endonuklease I-PpoI
6. SEQ ID NO: 6 Oligonukleotid-Primer p19
5'-TAAGGCCCTCGGTAGCAACGG-3'
7. SEQ ID NO: 7 Oligonukleotid-Primer p20
5'-GGGGTACCAAATCCAACCTAG-3'
8. SEQ ID NO: 8 Oligonukleotid-Primer p21:
5'-GGAGCTCGCTCCCCCGCCGTCGTTC-3'
9. SEQ ID NO: 9 Oligonukleotid-Primer p22
5'-GATGCATGATGACTTGACGGCATCCTC-3'
10. SEQ ID NO: 10
Nukleinsäuresequenz kodierend für das Transitpeptid der
kleine Untereinheit (SSU) der Ribulosebisphosphatcarboxylase
(Rubisco ssu) aus Erbse
11. SEQ ID NO: 11
Transitpeptid der kleine Untereinheit (SSU) der Ribulose-
bisphosphatcarboxylase (Rubisco ssu) aus Erbse
12. SEQ ID NO: 12
Transitpeptid der plastidären Transketolase aus Tabak.
13. SEQ ID NO: 13
Nukleinsäuresequenz kodierend für das Transitpeptid der
plastidären Transketolase aus Tabak (Leseraster 1; pTP09)
14. SEQ ID NO: 14
Nukleinsäuresequenz kodierend für das Transitpeptid der
plastidären Transketolase aus Tabak (Leseraster 2; pTP10)
15. SEQ ID NO: 15
Nukleinsäuresequenz kodierend für das Transitpeptid der
plastidären Transketolase aus Tabak (Leseraster 3; pTP11)
16. SEQ ID NO: 16
Transitpeptid der plastidären Isopentenylpyrophosphat
Isomerase-2 (IPP-2) aus *Arabidopsis thaliana*.

17. SEQ ID NO: 17
Nukleinsäuresequenz kodierend für das Transitpeptid der
plastidären Isopentenylpyrophosphat Isomerase-2 (IPP-2)
aus *Arabidopsis thaliana* (Leseraster 1; IPP-9)
- 5
18. SEQ ID NO: 18
Nukleinsäuresequenz kodierend für das Transitpeptid der
plastidären Isopentenylpyrophosphat Isomerase-2 (IPP-2)
aus *Arabidopsis thaliana* (Leseraster 2; IPP-10)
- 10
19. SEQ ID NO: 19
Nukleinsäuresequenz kodierend für das Transitpeptid der
plastidären Isopentenylpyrophosphat Isomerase-2 (IPP-2)
aus *Arabidopsis thaliana* (Leseraster 3; IPP-11)
- 15
20. SEQ ID NO: 20
Nukleinsäuresequenz kodierend für den PrbcL Promotor aus
Tabak.
- 20
21. SEQ ID NO: 21
Nukleinsäuresequenz kodierend für den Prps16-107 Promotor
aus Tabak.
22. SEQ ID NO: 22
Nukleinsäuresequenz kodierend für den Prn16 Promotor aus
Tabak.
- 25
23. SEQ ID NO: 23
Nukleinsäuresequenz kodierend für Promotor PaccD-129 aus
Tabak.
- 30
24. SEQ ID NO: 24
Nukleinsäuresequenz kodierend für den Promotor PclpP-53 aus
Tabak.
- 35
25. SEQ ID NO: 25
Nukleinsäuresequenz kodierend für den Promotor Prn-62 aus
Tabak.
- 40
26. SEQ ID NO: 26
Nukleinsäuresequenz kodierend für den Promotor Prps16 aus
Tabak.
- 45
27. SEQ ID NO: 27
Nukleinsäuresequenz kodierend für den Promotor PatpB/E-290
aus Tabak.

28. SEQ ID NO: 28
Nukleinsäuresequenz kodierend für den Promotor PrpoB-345 aus Tabak.
- 5 29. SEQ ID NO: 29
Nukleinsäuresequenz kodierend für eine Sequenz abgeleitet von der Konsensussequenz der $\sigma 70$ Promotoren aus E.coli.
30. SEQ ID NO: 30
10 Nukleinsäuresequenz kodierend für die 5'-untranslatierte Region des psbA Gens aus Tabak (5'psbA)
31. SEQ ID NO: 31
15 Nukleinsäuresequenz kodierend für die 5'-untranslatierte Region einschließlich 5' Anteilen aus dem kodierenden Bereich des rbcL Gens aus Tabak (5'rbcL).
32. SEQ ID NO: 32
20 Nukleinsäuresequenz kodierend für die 5'-untranslatierte Region des rbcLs Gens aus Tabak.
33. SEQ ID NO: 33
25 Nukleinsäuresequenz kodierend für die 3'-untranslatierte Region des psbA-1 Gens aus Synechocystis (3'psbA-1)
34. SEQ ID NO: 34
Nukleinsäuresequenz kodierend für die 3'-untranslatierte Region des psbA Gens aus Tabak (3'psbA)
- 30 35. SEQ ID NO: 35
Nukleinsäuresequenz kodierend für die 3'-untranslatierte Region des rbcL Gens aus Tabak (3'rbcL)
36. SEQ ID NO: 36
35 Nukleinsäuresequenz kodierend für synthetische Ribosomenbindestellen (RBS)
37. SEQ ID NO: 37
40 Nukleinsäuresequenz kodierend für synthetische Ribosomenbindestelle (RBS)
38. SEQ ID NO: 38 Oligonukleotid-Primer p65:
5'-ATGGATCCATATGGCCATGGCACAAGGGGTTGTG-3'
- 45 39. SEQ ID NO: 39 Oligonukleotid-Primer p66:
5'-CCCGGGCTCGAGCTGCAGCTACGCCGCTACGTC -3'

40. SEQ ID NO: 40
Insert von Vektor pCB193-17 umfassend Expressionskassette
kodierend für Fusionsprotein aus PLS und Φ C31 Rekombinase
- 5 41. SEQ ID NO: 41
Aminosäuresequenz des Fusionsproteins aus PLS und Φ C31
Rekombinase
42. SEQ ID NO: 42
10 Aminosäuresequenz der Φ C31 Rekombinase
43. SEQ ID NO: 43 Oligonukleotid-Primer p71:
5'-TGGATCCATATGGCCATGGCTAAGAAAGTAG-3'
- 15 44. SEQ ID NO: 44 Oligonukleotid-Primer p72:
5'-CCTCGAGCTGCAGTTAAGCGAGTTGG-3'
45. SEQ ID NO: 45
Binärvektor pCB245-37 umfassend Expressionskassette kodierend
20 für Fusionsprotein aus PLS und TP901-1 Rekombinase
46. SEQ ID NO: 46
Aminosäuresequenz des Fusionsproteins aus PLS und TP901-1
Rekombinase
- 25 47. SEQ ID NO: 47
Aminosäuresequenz der TP901-1 Rekombinase
48. SEQ ID NO: 48
30 Vektor pCB249-24 umfassend attP Erkennungsregion für die
Rekombinase aus TP901-1 und attP Erkennungsregion für die
Rekombinase aus Φ C31
49. SEQ ID NO: 49
35 Vektor pCB384-40 umfassend attP Erkennungsregion für die
Rekombinase aus TP901-1 und attP Erkennungsregion für die
Rekombinase aus Φ C31 sowie eine Expressionskassette für die
Homing Endonuklease I-PpoI
- 40 50. SEQ ID NO: 50 Nukleinsäuresequenz kodierend für das Insert
aus Vektor pCB456-2
51. SEQ ID NO: 51 Oligonukleotidprimer p354
45 5'-CGGATCCGGAGGCATCATGACTAAGAAAGTAGCAATCTATACACG

81

52. SEQ ID NO: 52 Oligonukleotidprimer p355
5'-ATAAGCTTTATTCTCGATATCG-3'
- 5 53. SEQ ID NO: 53 Nukleinsäuresequenz kodierend für das Insert
aus Vektor pCB557-1
54. SEQ ID NO: 54 Nukleinsäuresequenz kodierend für das Insert
aus Vektor pCB554-3
- 10 55. SEQ ID NO: 55 Oligonukleotidprimer p268
5'-TCGACTTGACATTCACTCTTCAATTATCTATAATGATACATG-3'
56. SEQ ID NO: 56 Oligonukleotidprimer p72
5'-CCTCGAGCTGCAGTTAAGCGAGTTGG-3'
- 15 57. SEQ ID NO: 57 Nukleinsäuresequenz kodierend für PCR Produkt
TP-Rek
58. SEQ ID NO: 58 Nukleinsäuresequenz kodierend für Southern-
20 Blot Sonde
59. SEQ ID NO: 59 Nukleinsäuresequenz kodierend für eine
Promotorsequenz abgeleitet von der
Konsensussequenz der $\sigma 70$ Promotoren
25 aus E.coli.
60. SEQ ID NO: 60 Nukleinsäuresequenz kodierend für das Insert
aus Vektor pCB487-20 umfassend Sequenz
kodierend für TP901 Rekombinase
- 30 61. SEQ ID NO: 61 Aminosäuresequenz kodierend für TP901
Rekombinase

Abbildungen

- 35 Im Rahmen des erfindungsgemäßen Verfahrens sind insbesondere die
in nachfolgenden Abbildungen verdeutlichten Ausführungsformen
besonders bevorzugt. Für die Abbildungen gelten allgemein nach-
folgende Abkürzungen:

40

X/Y : Rekombinase Erkennungsstelle

U/W: Rekombinase Erkennungsstelle, kann bevorzugt nicht mit
X/Y rekombinieren

45

- X'/Y': Rekombinase Erkennungsstelle, kann bevorzugt nur mit X/Y rekombinieren
- 5 U'/W': Rekombinase Erkennungsstelle, kann bevorzugt nur mit U/W rekombinieren
- 10 X/Y' bzw. X'/Y: Nach Rekombination zwischen X/Y und X'/Y' hervorgegangen, zwischen diesen Stellen kann die Rekombinase allein bevorzugt keine weiteren Rekombinationen katalysieren
- 15 U/W' bzw. U'/W: Nach Rekombination zwischen U/W und U'/W' hervorgegangen, zwischen diesen Stellen kann die Rekombinase allein bevorzugt keine weiteren Rekombinationen katalysieren
- R, R1, R2: Rekombinase
- 20 A, A' Paar homologer Sequenzen A und A'
- A/A' Ergebnis einer homologen Rekombination zwischen A und A' und/oder eines durch Reperatursynthese bedingten Austausches zwischen A und A'.
- 25 B, B' Paar homologer Sequenzen B und B'
- B/B' Ergebnis einer homologen Rekombination zwischen B und B' und/oder eines durch Reperatursynthese bedingten Austausches zwischen B und B'.
- 30 H1, H2: Paar homologer Sequenzen H1 und H2
- H1/2: Sequenz als Ergebnis der homologen Rekombination aus H1 und H2
- 35 DS funktionelle DSB-Erkennungssequenz
- nDS nicht funktionelle "Halbseite" einer DSB-Erkennungssequenz
- 40 E: DSBI-Enzym
- P, P': Promotor
- 45 I: weitere Nukleinsäuresequenz

S, S' Positive Selektionsmarker

NS Negativer Selektionsmarker

- 5 1. Fig.1A/B: Einführung einer RE-Sequenz in das Plastom mittels
Doppeltem "Cross-Over"

In einer besonders bevorzugten Ausführungsform 1 wird zunächst ein RE-Konstrukt in Plastiden einer höheren Pflanze eingebracht.

- 10 Das RE-Konstrukt wird in dieser Ausführungsform bevorzugt mit homologen Zielregionen und mit einem exprimierbaren Selektionsmarker (Promotor - 5'UTR - Selektionsmarker - 3'UTR) ausgerüstet und enthält in dieser Ausführungsform bevorzugt eine RE-Sequenz. Das RE-Konstrukt kann - optional - bereits weitere Gene von
15 Interesse kodieren. Es werden überwiegend homoplastome Masterpflanzen erzeugt.

Wie bereits oben beschrieben stellt A/A' bzw. B/B' das Ergebnis einer homologen Rekombination dar. Auch wenn diese Sequenzen in
20 nachfolgenden Schritten noch im Plastom vorhanden sind, sind sie in den entsprechenden Abbildungen nicht mehr dargestellt, da sie funktionell ohne Relevanz sind.

2. Fig.2: Einführung einer Insertionssequenz mit einer
25 Expressionskassette für eine Rekombinase und ggf. Selektionsmarker sowie weiteren Genen von Interesse

- Explantate der durch Ausführungsform 1 hergestellten Masterpflanzen werden für eine weitere Transformation mit einem
30 erfindungsgemäßen zirkulären Transformationskonstrukt genutzt. Die Rekombinase wird dabei von dem Transformationskonstrukt selber unter Kontrolle eines in Plastiden funktionellen Promotors exprimiert.

- 35 3. Fig.3: Einführung einer Insertionssequenz mit ggf. Selektionsmarker sowie weiteren Genen von Interesse und Expression/Einbringung der Rekombinase in trans

- In weiteren bevorzugten Ausführungsform wird die Rekombinase
40 nicht von dem Transformationskonstrukt kodiert, sondern entweder in trans exprimiert (in Plastiden oder als PLS-Fusionsprotein im Kern) bzw. in Form von RNA oder als Protein in die Plastiden transfiziert. Diese Ausführungsform ist insbesondere dann bevorzugt, wenn das Transformationskonstrukt keine Promotorelemente
45 umfasst, und eine Expression der kodierten Gene erst nach Inser-

tion in das Plastom unter Verwendung plastidärer, endogener Promotoren realisiert wird.

4. Fig.4: Einführung eines RE-Konstruktes mit zwei RE-Sequenzen

5

In einer weiteren Ausführungsform 2 wird zunächst ein RE-Konstrukt mit zwei RE-Sequenzen in Plastiden einer höheren Pflanze eingebracht. Das RE-Konstrukt wird in dieser Ausführungsform bevorzugt mit homologen Zielregionen und mit einem exprimierbaren Selektionsmarker (Promotor - 5'UTR - Selektionsmarker - 3'UTR) ausgerüstet. Dabei können die beiden eingebrachten RE-Sequenzen gleich oder unterschiedlich sein, können aber bevorzugt nicht - unter Einfluss einer Rekombinase - miteinander rekombinieren.

15

Das RE-Konstrukt kann - optional - bereits weitere Gene von Interesse kodieren. Es werden überwiegend homoplastome Masterpflanzen erzeugt.

20 5. Fig.5: Einführung einer Insertionssequenz mit zwei RE-Sequenzen und mit ggf. Selektionsmarker sowie weiteren Genen von Interesse

Explantate der unter Ausführungsform 2 (vgl. Fig. 4) hergestellten Masterpflanzen werden für eine weitere Transformation mit einem erfindungsgemäßen Transformationskonstrukt genutzt, das wiederum zwei RE-Sequenzen hat. Das Konstrukt kann linear oder zirkular sein (Fig. 5). Dabei können die beiden RE-Sequenzen gleich oder unterschiedlich sein, können aber bevorzugt nicht - unter Einfluss einer Rekombinase - miteinander rekombinieren.

Die Rekombinase bzw. die Rekombinasen können dabei von dem Transformationskonstrukt selber unter Kontrolle eines in Plastiden funktionellen Promotors exprimiert oder aber auch in trans exprimiert werden (in Plastiden oder als PLS-Fusionsprotein im Kern) bzw. in Form von RNA oder als Protein in die Plastiden transfiziert werden.

40 6. Fig.6: Einführung eines RE/DSB-Konstruktes mit einer RE-Sequenz und einer DSB-Erkennungssequenz

In einer weiteren Ausführungsform 3 wird zunächst ein RE/DSB-Konstrukt mit mindestens einer RE-Sequenz und einer DSB-Erkennungssequenz in Plastiden einer höheren Pflanze eingebracht. Das RE-Konstrukt wird in dieser Ausführungsform bevorzugt mit homologen

Zielregionen und mit einem exprimierbaren Selektionsmarker (Promotor - 5'UTR - Selektionsmarker - 3'UTR) ausgerüstet.

Das RE-Konstrukt kann - optional - bereits weitere Gene von Interesse - wie beispielsweise einen negativen Selektionsmarker - kodieren. Es werden überwiegend homoplastome Masterpflanzen erzeugt.

7. Fig.7A-C: Einführung einer Insertionssequenz sowie Expression/
10 Einbringung einer Rekombinase und eines DSBI-Enzyms

Explantate der gemäß Ausführungsform 3 hergestellten Masterpflanzen (vgl. Fig. 6) werden für eine weitere Transformation mit einem erfindungsgemäßen Transformationskonstrukt genutzt, das
15 wiederum eine RE-Sequenz hat und zusätzlich Sequenzabschnitte enthält, die zu denen im RE/DSB-Konstrukt homolog sind (z.B. Sequenz eines Selektionsmarkers oder eines teils derselben). Durch Wirkung einer Rekombinase entsteht ein entsprechendes Rekombinationsprodukt (Fig. 7A).

20 Die Rekombinase kann dabei von dem Transformationskonstrukt selber unter Kontrolle eines in Plastiden funktionellen Promotors exprimiert oder aber auch in trans exprimiert werden (in Plastiden oder als PLS-Fusionsprotein im Kern) bzw. in Form
25 von RNA oder als Protein in die Plastiden transfiziert werden.

Das Rekombinationsprodukt wird dann der Wirkung eines DSBI-Enzym ausgesetzt, dass die durch das RE/DSB-Konstrukt eingeführte DSB-Erkennungssequenz erkennt und schneidet. Das DSBI-Enzym kann da-
30 bei von dem Transformationskonstrukt selber unter Kontrolle eines in Plastiden funktionellen Promotors exprimiert oder aber auch in trans exprimiert werden (in Plastiden oder als PLS-Fusionsprotein im Kern) bzw. in Form von RNA oder als Protein in die Plastiden transfiziert werden. Durch den Schnitt wird eine intramolekulare
35 homologe Rekombination zwischen homologen Sequenzen erzeugt. In dem dargestellten Fall werden diese homologen Sequenzen beispielsweise durch eine Duplikation eines Selektionsmarkers oder Teile desselben dargestellt. Die homologe Rekombination bewirkt die Deletion der DSB-Erkennungssequenz. In einer bevorzugten Aus-
40 führungsform wird durch die homologe Rekombination zusätzlich ein negativer Selektionsmarker deletiert (Fig. 7B).

Durch das weiterhin vorliegende DSBI-Enzym kann die so erzeugte Plastomart nicht mehr geschnitten werden, während die Kopien
45 der mit dem RE/DSB-Konstrukt erzeugten Masterpflanze, die noch nicht die Insertionssequenz durch Rekombination inseriert haben, nach wie vor geschnitten werden können. Dadurch baut sich ein

Selektionsdruck auf, der - unterstützt durch eine Reparatursynthese - zu einer schnellen Anreicherung der erwünschten plastidären DNA-Moleküle führt (Fig. 7C).

- 5 Dem Fachmann ist bewusst, dass die Reihenfolge der exprimierten Gene in einem Operon austauschbar ist und insofern in den oben beschriebenen Ausführungsformen variieren kann. Prinzipiell kann die Rekombinase und/oder das DSBI-Enzym auf dem Transformationskonstrukt und/oder separat (im Kern oder Plastiden) exprimiert
10 oder andersweitig - beispielsweise durch Transfektion mit RNA oder Protein in Plastiden - eingebracht werden.

8. Fig.8: Deletion von Sequenzen mittels intramolekularer homologer Rekombination induziert durch sequenzspezifische Doppelstrangbrüche
15

- Sequenzen - beispielsweise kodierend für Selektionsmarker bzw. DSBI-Enzyme - sind bei allen oben beschriebenen Ausführungsformen bevorzugt von Homologiesequenzen H1 und H2 flankiert, die
20 eine ausreichende Länge und Homologie aufweisen, um miteinander zu rekombinieren. Die Rekombination wird durch die Induktion mindestens einen Doppelstrangbruches zwischen den beiden Homologiesequenzen gelegenen DSB-Erkennungssequenz induziert. Zur Induktion des Doppelstrangbruches wird bevorzugt ein DSBI-Enzym
25 transient exprimiert oder eingebracht (Fig. 8).

- Dem Fachmann ist bewusst, dass die Reihenfolge der exprimierten Gene in einem Operon austauschbar ist und insofern in den oben beschriebenen Ausführungsformen variieren kann. Auch
30 kann bei Verwendung nur einer Homologiesequenz zu Insertion der Insertionssequenz diese an der 5'- (wie in den Abbildungen gezeigt) oder 3'-Seite des Doppelstrangbruches lokalisiert sein. Prinzipiell kann das DSBI-Enzym auf dem Transformationskonstrukt und/oder separat (im Kern oder Plastiden) exprimiert oder anders-
35 weitig - beispielsweise durch Transfektion mit RNA oder Protein in Plastiden - eingebracht werden.

9. Fig.9: Southern Analyse von überwiegend homotransplastomen Pflanzen
40

- Wildtyp und überwiegend homotransplastome Masterpflanzen wurden bezüglich der Modifikation (Einführung einer RE-Sequenz) analysiert (vgl. Beispiel 4). Durch die Modifikation wurde eine Bande von 1750 bp detektiert (Bahn 2, 3, und 4 entsprechend den Linien
45 CB199NTH-4, -6, und -8), während in der nicht modifizierten Wildtyppflanze eine Bande von 3100 bp detektiert wurde (Bahn 1).

Fig. 10: Southern Analyse von überwiegend homotransplastomen Pflanzen

- A: Wildtyp und überwiegend homotransplastome Masterpflanzen wurden bezüglich der Modifikation (Einführung einer RE- und DSB-Erkennungssequenz) in einem Southern-Blot analysiert (vgl. Beispiel 4). Durch die Modifikation wurde in der hier mit HindIII behandelten DNA eine Bande von etwa 3,8 kb detektiert (Bahn 2, 3 und 4 entsprechend den Linien CB199NTH-4, -6, und -8), während in der nicht modifizierten Wildtyp-pflanze eine Bande von etwa 7,7 kb detektiert wurde (Bahn 5). In den Linien CB199NTH-4, -6, und -8 ist neben der erwarteten etwa 3,8 kb Bande und der etwa 7,7 kb Bande noch eine Bande bei etwa 5,2 kb zu erkennen. Diese Bande ergibt sich (wie in B verdeutlicht) aus der Hybridisierung des 16SrRNA-Promotors, welcher in der Sonde enthalten ist, mit dem 16SrRNA Promotor, der stromabwärts des Selektionsmarker aadA zu finden ist. Diese etwa 5,2 kb große Bande ist wesentlich schwächer als die 3,8 kb große Bande weil nur ein kleiner Anteil der Sonde tatsächlich mit diesem Fragment hybridisiert (wt, nicht modifizierte Wildtyppflanze; m - Molekularer Größenstandard)
- B: Schematische Darstellung der HindIII Fragmente, welche in A durch die Hybridisierung mit der Sonde (probe) in einer nicht modifizierte Wildtyppflanze (oben) und einer modifizierten Pflanze CB199NTH (unten) zu erwarten waren. (trnV - Gen codierend für eine tRNA-Val; rrn16 - Gen codierend für die 16SrRNA; aadA - Gen codierend für einen Selektionsmarker; 3'psbA - nicht kodierender Bereich stromabwärts des psbA Gens, hier eingebaut in die Expressionskassette für den Selektionsmarker aadA)

Fig. 11: Southern Analyse von überwiegend homotransplastomen Pflanzen

- A: Wildtyp und überwiegend homotransplastome Masterpflanzen wurden bezüglich der Modifikation (Einführung einer RE- und DSB-Erkennungssequenz) in einem Southern analysiert (vgl. Beispiel 14). Durch die Modifikation wurde in der hier mit EcoRI behandelten DNA eine Bande von etwa 1,7 kb detektiert (Bahn 1 und 4 entsprechend den Linien CB456NTH-1 und -15), während in der nicht modifizierten Wildtyppflanze eine Bande von etwa 3,1 kb detektiert wurde (Bahn 6). (wt - nicht modifizierte Wildtyppflanze; wildtype - zeigt die erwartete Fragmentgröße in nicht modifizierten Wildtyppflanzen an; trans-

genic - zeigt die erwartete Fragmentgröße in Pflanzen
CB456NTH an)

- 5 B: Schematische Darstellung des EcoRI Fragments, welches in A
durch die Hybridisierung mit der Sonde (probe) in einer modi-
fizierten Pflanze CB456NTH zu erwarten war. (trnV - Gen
codierend für eine tRNA-Val; rrn16 - Gen codierend für die
16SrRNA; aadA - Gen codierend für einen Selektionsmarker;
10 3'psbA (Synec) - nicht kodierender Bereich stromabwärts
des psbA-1 Gens aus Synechocystis, hier eingebaut in die
Expressionskassette für den Selektionsmarker aadA; Psynth. -
synthetischer Promotor abgeleitet von der Konsensussequenz
für E. coli σ 70 Promotoren).

15 Beispiele

Allgemeine Methoden:

- Die chemische Synthese von Oligonukleotiden kann beispielsweise,
20 in bekannter Weise, nach der Phosphoramiditmethode (Voet, Voet,
2. Auflage, Wiley Press New York, Seite 896-897) erfolgen. Die
im Rahmen der vorliegenden Erfindung durchgeführten Klonierungs-
schritte wie z.B. Restriktionsspaltungen, Agorosegelelektro-
phorese, Reinigung von DNA-Fragmenten, Transfer von Nuklein-
25 säuren auf Nitrozellulose und Nylonmembranen, Verknüpfen von
DNA-Fragmenten, Transformation von E. coli Zellen, Anzucht von
Bakterien, Vermehrung von Phagen und Sequenzanalyse rekombinanter
DNA werden wie bei Sambrook et al. (1989) Cold Spring Harbor
Laboratory Press; ISBN 0-87969-309-6 beschrieben durchgeführt.
30 Die Sequenzierung rekombinanter DNA-Moleküle erfolgt mit einem
Laserfluoreszenz-DNA-Sequenzierer ALF-Express (Pharmacia, Upsala,
Schweden) nach der Methode von Sanger (Sanger et al. (1977) Proc
Natl Acad Sci USA 74:5463-5467).

35 Beispiel 1: Erstellen eines Basis Vektors für die Plastidentrans- formation

Zunächst wurden aus dem Plastom der Tabak Varietät SR1 die aus-
gewählte Zielregionen mittels PCR kloniert. Die linke Zielregion
40 wurde dabei mit den Primern p19 und p20 amplifiziert.

p19: 5'-TAAGGCCCTCGGTAGCAACGG-3' (SEQ ID NO: 6)

p20: 5'-GGGGTACCAAATCCAACACTAG-3' (SEQ ID NO: 7)

Zur Amplifikation der rechten Zielregion wurden die Primer p21 und p22 verwendet, wobei durch letzteren Primer in den amplifizierten Teil der 16SrDNA zusätzlich zur SR1 Resistenz (binding type marker) auch noch eine Spectinomycin-Resistenz eingebracht wurde.

p21: 5'-GGAGCTCGCTCCCCGCGCTCGTTC-3' (SEQ ID NO: 8)

p22: 5'-GATGCATGATGACTTGACGGCATCCTC-3' (SEQ ID NO: 9)

10

Die beiden amplifizierten Regionen wurden in pBluescript bzw. pZeroBlunt kloniert und sequenziert. Die linke und rechte Zielregionen wurden anschließend in das Rückgrat des pUC19 Vektors kloniert. Dazu wurden die Schnittstellen Eco0109I und PvuII des Vektors genutzt. Zwischen die linke und rechte Zielregion wurde eine multiple Klonierungsstelle aus dem pBluescript (von KpnI bis SacI) kloniert. Diese befindet sich im Basisvektor für die Plastidentransformation zwischen den beiden plastidär kodierten Genen trnV und rrn16. Dieser Basisvektor für die Plastidentransformation erhielt die Bezeichnung pCB42-94 (SEQ ID NO: 1). Der Vektor enthält nachfolgende Sequenzelemente:

- a) Position komplementär bp 55-1405: Rechte Zielregion mit dem partiellen Gen der 16SrRNA (komplementär bp 56 bis 1322). In letzterem sind Mutationen für die Streptomycin-Resistenz (SR1, Position bp 346) und Spectinomycin-Resistenz (SPC1, Position bp 68).
- b) Position komplementär bp 2374 bis 1510: Linke Zielregion umfassend u.a. ORF131 (bp 1729 bis 2124) und trnV Gen (komplementär bp 1613 bis 1542).
- c) Position bp 1404 bis 1511: multiple Klonierungsstelle
- d) Position bp 2629 bis 3417: Ampicilin Resistenz im Vektor-rückgrat

Beispiel 2: Erstellen eines Vektors (pCB199-3) für die Einbringung einer nicht natürlicherweise vorkommenden DSB-Erkennungssequenz für die Homing Endonukleasen I-PpoI und RE-Sequenzen (attB) für die Rekombinasen Φ C31 und TP901-1 in das Plastom von Tabak

In die multiple Klonierungsstelle des Basisvektors pCB42-94 (SEQ ID NO: 1) für die Plastidentransformation wurden sukzessive verschiedene Elemente kloniert:

- a) frt Erkennungsregion (mutiert, enthält keine XbaI Schnittstelle; komplementär 1307-1354)
- b) Expressionskassette zur Expression des Markergens aadA
5 bestehend aus:
 - i) Promotor des Gens für 16SrRNA (komplementär 1191-1281)
 - 10 ii) 5' untranslatierter Bereich des Tabak rbcL Gens (komplementär 1167-1184) einschließlich mutierter 5' Anteile des rbcL Gens (Duplikation von 6 AS des rbcL Gens, teilweise mutiert, als Folge der Klonierungsstrategie, so dass eine Fusion codierend für insgesamt 12 Aminosäuren (komplementär 1131-1166) mit dem nachfolgendem Element, 15 dem aadA Gen, entstand)
 - iii) aadA Gen (komplementär 336-1130)
 - iv) der 3' Bereich des psbA Gens (komplementär 232-323)
 - 20 c) attB RE-Sequenzen für die Rekombinasen Φ C31 (Position Basenpaar 131 bis 170) und TP901-1 (Position Basenpaar 21 bis 121)
 - d) Core-Erkennungsregion für die Homing Endonuklease I-PpoI
25 (Position Basenpaar komplementär 176 bis 190).

Anstelle der multiplen Klonierungsstelle im Basisvektor für die Plastidentransformation enthält dieser Vektor mit der Bezeichnung pCB199-3 die aufgeführten Elemente im Rahmen der Nukleinsäure-
30 sequenz mit der SEQ ID NO: 2. Angegeben ist die Sequenz, die die komplette MCS von KpnI bis SacI ersetzt. In der angegebenen Sequenz gibt es aufgrund der Klonierungsstrategie jedoch keine KpnI Schnittstelle mehr.

35 Beispiel 3: Erstellen eines weiteren Vektors (pCB401-20) für die Einbringung einer nicht natürlicherweise vorkommende Erkennungsregion für die Homing Endonukleasen I-PpoI und RE-Sequenzen (attB) für die Rekombinasen Φ C31 und TP901-1 in das Plastom von Tabak

40 Im Gegensatz zu dem in Beispiel 2 beschriebenen Vektor pCB199-3 enthält der hier beschriebene Vektor keinen Promotor und 3'UTR, der direkt mit dem Selektionsmarker aadA verknüpft wurde. Vielmehr wird die Expression des aadA Gens ausgehend von dem Promotor
45 des trnV Gens, welches im Plastidengenom bzw. in der linken Zielregion stromaufwärts des aadA Gens lokalisiert ist, gesteuert. Ziel bei der Erstellung dieses Vektors war es, Sequenzdupli-

91

kationen durch das Ausnutzen von regulatorischen Bereichen aus dem Tabakplastidengenom zu vermeiden. Dazu wurden in die multiple Klonierungsstelle des Basisvektors pCB42-94 für die Plastidentransformation sukzessive verschiedene Elemente kloniert:

5

a) Ribosomenbindestelle (komplementär bp 1033 bis 1050)

b) aadA Gen (komplementär bp 238 bis 1032)

10 c) attB RE-Sequenzen für die Rekombinasen Φ C31 (Position Basenpaar 131 bis 170) und TP901-1 (Position Basenpaar 21 bis 121)

d) Core-Erkennungsregion für die Homing Endonukleasen I-PpoI (komplementär bp 176 bis 190)

15

Der so erhaltene Vektor vermittelt auch in *E. coli* eine Spectinomycin Resistenz. Anstelle der multiplen Klonierungsstelle im Basisvektor pCB42-94 (SEQ ID NO: 1) für die Plastidentransformation enthält dieser Vektor mit der Bezeichnung pCB401-20

20 die aufgeführten Elemente im Rahmen der Nukleinsäuresequenz mit der SEQ ID NO: 3. Auch hier ist die gesamte die MCS ersetzende Sequenz (von SacI bis KpnI) angegeben.

Beispiel 4: Erstellen von überwiegend homoplastomischen Tabak
25 Masterpflanzen, die eine nicht natürliche DSB-Erkennungssequenz und RE-Sequenzen (attB) enthalten

Das Plasmid pCB199-3 wurde in die Plastiden von Tabak (*Nicotiana tabacum* cv. Petit Havana) wie im folgenden beschrieben einge-
30 bracht. Die regenerierten Pflanzen wurden CB199NTH genannt. Unabhängige Linien wurden mit verschiedenen Endnummern (zum Beispiel CB199NTH-4) versehen.

Analog wird der Vektor pCB401-20 in die Plastiden von Tabak ein-
35 gebracht. Die resultierenden Pflanzen werden entsprechend mit CB401NTH bezeichnet.

Zunächst wurden von in vitro gewachsenen Pflanzen mit einem sterilen Korkbohrer Blattscheibchen mit einem Durchmesser von
40 2,0 bis 2,5 cm ausgestochen und je mit der Blattoberseite auf eine Petrischale mit Beschussmedium (MS-Salze (Sigma-Aldrich): 4,3 g/L; Saccharose: 30,0 g/L, Phytoagar (Duchefa, P1003): 0,6 % (w/v); pH 5,8. Nach dem Autoklavieren wurden 1,0 mg/L Thiamin (Duchefa, T0614) und 0,1 g/L myo-Inositol (Duchefa, I0609) zuge-
45 geben). Die vom Agar abgewandte Blattunterseite wurde mittels der Partikelkanone anschließend beschossen. Dazu wurde zunächst die zu transformierende Plasmid-DNA (aus *E. coli* gereinigt mittels

Nucleobond AX100 Macherey & Nagel) nach folgendem Protokoll auf 0,6 µm große Goldpartikel aufgebracht ("Coating"). Zunächst wurden 30 mg Goldpulver (BioRad) in Ethanol aufgenommen. 60 µl von der Goldsuspension werden in ein neues Eppendorfggefäß über-
5 führt und die Goldpartikel durch Zentrifugation (für 10 Sekunden) sedimentiert. Die Goldpartikel wurden zweimal in je 200 µl sterilem Wasser gewaschen und nach einem weiteren Zentrifugationsschritt in 55 µl Wasser aufgenommen. Unter ständigem Mischen (Vortexen) wurden schnell hinzugegeben:

10

- 5 µl Plasmid-DNA (1 µg/µl)
- 50 µl 2,5 M CaCl₂
- 20 µl 0,1 M Spermidine

15 Die Suspension wurde sodann für weitere 3 Minuten auf dem Vortex gemischt und anschließend kurz anzentrifugiert. Die sedimentierten Gold-DNA-Komplexe wurden 1 bis 2 mal in je 200 µl Ethanol gewaschen und nach einem weiteren Zentrifugationsschritt schließlich in 63 µl Ethanol aufgenommen. Pro Schuss wurden 3,5 µl (ent-
20 sprechend 100 µg Gold) dieser Suspension auf einen Macrocarrier aufgebracht.

Gemäß den Herstellerangaben wurde die Partikelkanone (BioRad, PDS1000He) vorbereitet und die Blattexplantate in einem Abstand
25 von 10 cm mit den Gold-DNA-Komplexen beschossen. Dabei wurden folgende Parameter verwendet: Vakuum: 27 inch Hg, Druck 1100 psi. Die Explantate werden nach dem Beschuss für 2 Tage in Klimakammern (24°C, 16 h Licht, 8 h Dunkelheit) inkubiert und anschließend mit einem Skalpell in etwa 0,5 cm² große Segmente zerteilt. Diese
30 Segmente wurden dann auf Regenerationsmedium [Beschussmedium zuzüglich 1 mg/L 6-Benzylaminopurin (BAP, Duchefa, B0904) und 0,1 mg/L Naphthyllessigsäure (NAA, Duchefa, N0903)] zuzüglich 500 mg/L Spectinomycin (Duchefa, S0188) überführt und für 10 bis 14 Tage unter oben genannten Bedingungen in einer Klimakammer inkubiert.
35 Nach Ablauf dieser Zeit wurden die Blattsegmente auf frisches Regenerationsmedium zuzüglich 500 mg/L Spectinomycin überführt. Dieser Vorgang wurde so lange wiederholt, bis sich grüne Sprosse an den Explantaten bildeten. Die Sprosse wurden mit einem Skalpell abgetrennt und auf Anzuchtmedium (wie Beschussmedium, jedoch
40 10 g/L Saccharose anstelle von 30 g/L Saccharose) zuzüglich 500 mg/L Spectinomycin angezogen.

Optional können, um möglichst überwiegend homoplastome Pflanzen zu erhalten, von den regenerierten Pflanzen wiederum Explantate
45 abgeschnitten und auf Regenerationsmedium mit 1000 mg/L Spectinomycin gelegt werden. Regenerierende Sprosse werden in Gläser mit Anzuchtmedium zuzüglich 500 bis 1000 mg/L Spectinomycin gesetzt.

93

Nach dem Bewurzeln werden die Pflanzen ins Gewächshaus transferiert und dort auf Erde bis zur Samenreife herangezogen.

Im Fall der Transformation mit dem Plasmid pCB199-3 wurden
5 8 Pflanzen erhalten, die Resistenz gegenüber Spectinomycin zeigten. Mittels PCR und Southern-Analysen wurde nachgewiesen, dass drei dieser Linien (Linien CB199NTH-4, -6 und -8) auch tatsächlich das aadA Gen in das Plastidengenom eingebaut haben.

- 10 Zur Analyse der transplastomen Pflanzen nach Southern wurde Gesamt-DNA aus Blättern von transformierten und nicht transformierten Pflanzen wurde mit Hilfe des GenElute Plant Genomic DNA Kit (Sigma) isoliert. Die DNA wurde dabei in 200 µl Eluat aufgenommen. 86 µl davon wurden je mit 10 µl 10x Restriktions-
- 15 puffer sowie 40 U Restriktionsendonuklease versetzt und 4 bis 8 h bei der für das Restriktionsenzym empfohlenen Temperatur inkubiert. Anschließend wurde die DNA in dem Fachmann bekannter Weise mit Ethanol präzipitiert und anschließend in 20 µl Wasser aufgenommen. Die Proben wurden anschließend gemäß dem Fachmann
- 20 bekannten Verfahren auf einem Agarosegel aufgetrennt, die DNA in dem Gel denaturiert und mittels Kapillarblot auf Nylonmembran übertragen.

- Eine entsprechende Sonde für die radioaktive Hybridisierung wurde
25 mit Hilfe des HighPrime (Roche) Systems erstellt. Die Membran wurde zunächst 1 h bei 65°C mit HybPuffer (1%(w/v) Rinderseumalbumin; 7%(w/v) SDS; 1 mM EDTA; 0,5 M Natriumphosphat-Puffer, pH 7,2) vorhybridisiert. Anschließend wurde die hitze-denaturierte Sonde hinzugegeben, und es wurde über Nacht bei 65°C
- 30 hybridisiert. Die Blots wurden anschließend wie folgt gewaschen: einmal Spülen mit 2xSSPE/0,1%SDS; für 15 min. bei 65°C waschen mit 2xSSPE/0,1%SDS; für 15 min. bei 65°C waschen mit 1xSSPE/0,1%SDS und ggf. wurde der letzte Schritt noch einmal wiederholt (20x SSPE ist 3 M NaCl; 0,2 M NaH₂PO₄; 0,5 M EDTA; pH 7,4).

35

Die Hybridisierung wurde anschließend mit Hilfe eines Phosphorimagers (Molecular Imager FX, BioRad) analysiert.

- Beispielsweise wurde mit PstI geschnittene Gesamt-DNA verschiedener Pflanzen, die nach der Transformation mit pCB199-3 regeneriert wurden, mit dem aadA Gen als radioaktiv markierte Sonde (793bp PstI\NcoI Fragment aus pCB199-3) hybridisiert. Dabei konnte festgestellt werden, dass die Linien CB199NTH-4, -6 und -8 tatsächlich das aadA Gen in die DNA eingebaut hatten. Des
- 40 weiteren wurde mit EcoRI und XhoI geschnittene gesamte DNA von CB199NTH-4, -6 und -8 mit einer radioaktiv markierten Sonde (1082 bp Bsp120I/SacI Fragment aus pCB199-3) hybridisiert, welche mit
- 45

einem Teil der 16S rDNA hybridisiert. Während im Wildtyp (nicht transformierte Pflanze) erwartungsgemäß eine Bande von etwa 3100 bp nachgewiesen wurde, wurde in den transplastomen Linien überwiegend eine Bande bei 1750 bp detektiert, die sich aufgrund
 5 der Insertion der Insertionssequenz aus pCB199-3 in das Plastom ergibt (Fig. 9). Die erhaltenen Pflanzen können als überwiegend homotransplastom verstanden werden.

Beispiel 5: Klonierung von Rekombinasen und des DSBI-Enzyms
 10 I-Ppo I

5.1 Klonieren der Rekombinase aus dem Phagen Φ C31 und nukleäre Transformation als PLS-Fusion

Die Rekombinase aus dem Phagen Φ C31 (DSM No. 49156, DSMZ-Deutsche
 15 Sammlung von Mikroorganismen und Zellkulturen GmbH, Braunschweig, Deutschland) wurde mittels PCR amplifiziert. Dazu wurden die folgenden Oligonukleotide benutzt:

20 p65: ATGGATCCATATGGCCATGGCACAAGGGGTTGTG (SEQ ID NO: 38)

p66: CCCGGGCTCGAGCTGCAGCTACGCCGCTACGTC (SEQ ID NO: 39)

Das erhaltene PCR Fragment wurde in pGEM-Teasy kloniert und sequenziert. Der resultierende Vektor wurde pCB109-28 benannt.

25 Anschließend wurde das für die Rekombinase kodierende Gen an Sequenzen fusioniert, die für das IPP-Transitpeptid codieren, und funktionell an den 35S Promotor und vor den ocs Terminator gekoppelt. Als Selektionsmarker war neben dem beschriebenen Element ein nos-Promotor - pat - nos-Terminator Element vor-
 30 handen, welches die Selektion transgener Pflanzen auf Phosphino-tricin erlaubte. Das fertige Binärkonstrukt wurde pCB193-17 genannt. Die Sequenz des Inserts ist unter SEQ ID NO: 40 beschrieben und enthält die nachfolgenden Elemente:

35 Position Basenpaar 7 bis 535: 35S Promotor

Position Basenpaar 563 bis 733: Sequenz kodierend für
 IPP-Transitpeptid

40 Position Basenpaar 749 bis 2590: Gen kodierend für die Φ C31
 Rekombinase

Position Basenpaar 2615 bis 2807: ocs Terminator

45 Das Plasmid pCB193-17 wurde mittels Agrobakterium-vermittelter Transformation in Tabak (N. tabacum cv. Petit Havana) nach dem Fachmann geläufige Verfahren transformiert. Transgene Pflanzen

aus dieser Transformation wurden mit CB193NTH bezeichnet. Die Expression der Φ C31 Rekombinase wurde mittels RealTime PCR nachgewiesen.

5 5.2 Klonieren der Rekombinase aus dem Phagen TP901-1 und nukleäre Transformation als PLS-Fusion

Die Rekombinase aus dem Phagen TP901-1 (lag als Prophage in einem lysogenen Stamm von *Lactococcus lactis* (Stamm 901-1) vor) wurde mittels PCR amplifiziert. Dazu wurden die folgenden Oligonukleotide benutzt:

p71: TGGATCCATATGGCCATGGCTAAGAAAGTAG (SEQ ID NO: 43)

15 p72: CCTCGAGCTGCAGTTAAGCGAGTTGG (SEQ ID NO: 44)

Das erhaltene PCR Fragment von 1486 bp wurde in pCR2.1-TOPO (Invitrogen) kloniert und sequenziert. Der resultierende Vektor wurde pCB127-2 genannt (zwei Unterschiede wurden in der Sequenz im Vergleich zur veröffentlichten Sequenz gefunden. Einer davon führt zu einem Aminosäureaustausch (V7A)). Anschließend wurde das für die Rekombinase kodierende Gen an Sequenzen fusioniert, die für das IPP-Transitpeptid codieren, und funktionell an den 35S Promotor und vor den ocs Terminator gekoppelt. Als Selektionsmarker war neben dem beschriebenen Element ein nos-Promotor - pat - nos-Terminator Element vorhanden, welches die Selektion transgener Pflanzen auf Phosphinotricin erlaubte. Das fertige Binärkonstrukt wurde pCB245-37 genannt. Die Sequenz des Inserts ist unter SEQ ID NO: 45 beschrieben und enthält die nachfolgenden Elemente:

Position Basenpaar 12 bis 540: 35S Promotor

35 Position Basenpaar 568 bis 738: Sequenz kodierend für IPP-Transitpeptid

Position Basenpaar 748 bis 2205: Gen kodierend für die TP901-1 Rekombinase

40 Position Basenpaar 2231 bis 2422: ocs Terminator

Das Plasmid pCB245-37 wurde mittels Agrobakterium-vermittelter Transformation in Tabak (*N. tabacum* cv. Petit Havana) nach dem Fachmann geläufige Verfahren transformiert. Transgene Pflanzen aus dieser Transformation wurden mit CB245NTH bezeichnet. Die

Expression der TP901-1 Rekombinase wurde mittels RealTime PCR nachgewiesen.

5.3 Klonieren der Homing Endonuklease I-PpoI

5

Die Homing Endonuklease I-PpoI wurde aus 26 synthetisch erstellten Oligonukleotiden mittels PCR in Anlehnung an die Methode von Stemmer WPC et al. (1995) Gene 164: 49-53 erstellt (SEQ ID NO: 4). Die zugrunde liegende Sequenz wurde von der veröffentlichten Sequenz abgeleitet (Accession No. M38131 Nukleotide 86 bis 577). Dabei wurden einige Mutationen eingeführt, um Erkennungsstellen für Restriktionsendonukleasen aus dem Gen zu entfernen, von denen aber keine eine veränderte Aminosäuresequenz bedingt.

15

Beispiel 6: Kreuzen von Masterpflanzen CB199NTH mit Rekombinase exprimierenden Pflanzen

6.1 Kreuzen von Masterpflanzen CB199NTH mit Φ C31 Rekombinase exprimierenden Pflanzen CB193NTH

20

Pollen der Pflanzen CB193NTH wurden auf die Narbe von Blüten der Pflanzen CB199NTH (Linien -4, -6 und -8) gestäubt. Zuvor wurden die Antheren der entsprechenden Blüten von den CB199NTH Pflanzen entfernt. Die Pflanzen wurden bis zur Samenreife weiter im Gewächshaus belassen. Die geernteten Samen wurden *in vitro* auf Anzuchtmedium (vgl. Beispiel 4; zuzgl. 500 mg/L Spectinomycin, 10 mg/L Phosphinotricin) ausgelegt und resistente Pflanzen unter sterilen Bedingungen weiter angezogen. Diese Pflanzen wurden mit CB199NTHxCB193NTH bezeichnet. Explantate von 4 bis 8 Wochen alten Pflanzen wurden für weitere Beschussexperimente (s.u.) genutzt.

25

30

6.2 Kreuzen von Masterpflanzen CB199NTH mit TP901-1 Rekombinase exprimierenden Pflanzen CB245NTH

35

Pollen der Pflanzen CB245NTH wurden auf die Narbe von Blüten der Pflanzen CB199NTH (Linien -4, -6 und -8) gestäubt. Zuvor wurden die Antheren der entsprechenden Blüten von den CB199NTH Pflanzen entfernt. Die Pflanzen wurden bis zur Samenreife weiter im Gewächshaus belassen. Die geernteten Samen wurden *in vitro* auf Anzuchtmedium (vgl. Beispiel 4; zuzgl. 500 mg/L Spectinomycin, 10 mg/L Phosphinotricin) ausgelegt und resistente Pflanzen unter sterilen Bedingungen weiter angezogen. Diese Pflanzen wurden mit CB199NTHxCB245NTH bezeichnet. Explantate von 4 bis 8 Wochen alten Pflanzen wurden für weitere Beschussexperimente (s.u.) genutzt.

40

45

97

Beispiel 7: Erstellen eines Vektors, der für die Integration mit Hilfe von Rekombinasen in das Plastom von entsprechenden Masterpflanzen geeignet ist

- 5 Sukzessive wurden die verschiedenen Elemente, die für so einen Vektor notwendig sind in das Rückgrat des pBLUESCRIPT kloniert. Die angegebene Sequenz ersetzt die multiple Klonierungsstelle des pBLUESCRIPT von KpnI bis SacI, wobei die KpnI Schnittstelle in der angegebenen Sequenz aufgrund der Klonierungsstrategie
- 10 nicht mehr funktionell vorhanden ist. Das Rückgrat des Vektors entspricht dem des pBLUESCRIPT. Der erhaltene Vektor wurde mit pCB249-24 bezeichnet. Das Insert ist durch SEQ ID NO: 48 beschrieben und umfasst nachfolgende Elemente:
- 15 Elemente:
- Position Basenpaar 861 bis 910: attP Erkennungsregion für die Rekombinase aus Φ C31
- 20 Position Basenpaar 925 bis 1031: Prps16 Promotor aus Tabak
- Position Basenpaar 1040 bis 1060: 5'rbcl (5' untranslatierter Bereich vom rbcL Gen aus Tabak)
- 25 Position Basenpaar 1076 bis 1879: nptII Gen
- Position Basenpaar 1916 bis 2053: 3'rbcl (3' untranslatierter Bereich vom rbcL Gen aus Tabak)
- 30 Beispiel 9: Erstellen eines Vektors, der für die Integration mit Hilfe von Rekombinasen in das Plastom von entsprechenden Masterpflanzen geeignet ist und einen Schnitt durch ein DSBI-Enzym ermöglicht
- 35 Sukzessive wurden die verschiedenen Elemente, die für einen solchen Vektor notwendig sind in das Rückgrat des pBLUESCRIPT kloniert. Die angegebene Sequenz ersetzt die multiple Klonierungsstelle des pBLUESCRIPT von KpnI bis SacI, wobei die KpnI Schnittstelle in der angegebenen Sequenz aufgrund der Klonierungsstelle
- 40 nicht mehr funktionell vorhanden ist. Das Rückgrat des Vektors entspricht dem des pBLUESCRIPT. Der erhaltene Vektor wurde mit pCB384-40 bezeichnet. Das Insert ist durch SEQ ID NO: 49 beschrieben und umfasst nachfolgende Elemente:

98

- i) Position Basenpaar 70 bis 765: homologe Region zu Plastomsequenzen in den Masterpflanzen CB199NTH - nach der Integration der hier betrachteten Insertionssequenz mittels PhiC31 Rekombinase liegt dieser Bereich dupliziert als direkter "repeat" (gleiche Orientierung) zu den homologen Sequenzen aus CB199NTH vor. Dies erlaubt die intrachromosomale Rekombination nach der Insertion der Insertionssequenz, was zur Eliminierung der I-PpoI Erkennungsregion aus den bereits durch Insertion modifizierten Plastomkopien führt. Darin enthalten komplementär von Basenpaar 70 bis 163 eine 3'psbA Sequenz und komplementär von Basenpaar 175 bis 765 ein Anteil vom aadA Gen.
- ii) Position Basenpaar 861 bis 910: attP Erkennungsregion für die Rekombinase aus PhiC31
- iii) Position Basenpaar 925 bis 1031: Prps16 Promotor aus Tabak
- iv) Position Basenpaar 1040 bis 1060: 5'rbcl (5' untranslatierter Bereich vom rbcl Gen aus Tabak)
- v) Position Basenpaar 1076 bis 1879: nptII Gen
- vi) Position Basenpaar 1897 bis 1974: 5'psbA (5' untranslatierter Bereich des psbA Gens aus Tabak)
- vii) Position Basenpaar 1975 bis 2466: codierende Region für die Homing Endonuklease I-PpoI
- viii) Position Basenpaar 2501 bis 2638: 3'rbcl (3' untranslatierter Bereich des rbcl Gens aus Tabak)

Beispiel 10: Transformation von CB199NTHxCB193NTH Pflanzen mit pCB249-24

- Zunächst wurde die Insertionssequenz in vitro aus dem Plasmid ausgeschnitten und diese anschließend zirkularisiert. 5 µg der DNA von pCB249-24 wurden dazu mit 10 U SalI (Roche) und 10U XhoI (Roche) in einem 20µl Restriktionsansatz unter den vom Hersteller angegebenen Bedingungen für 2h inkubiert. Anschließend wurde der Ansatz 20 Minuten bei 65°C inkubiert. Über Nacht wurde bei 14°C eine Ligationsreaktion durchgeführt (Zugabe von 5 µl 10x Ligasepuffer, 2U T4-Ligase (Roche) zum Restriktionsansatz und auffüllen auf 50µl mit Wasser). Unter den gewählten Bedingungen wurde überwiegend eine Rezirkularisierung der Fragmente beobachtet (nicht gezeigt). Die DNA wurde aus dem Ligationsansatz mit Hilfe des Wizard Plus Minipreps DNA Purification System (Promega) auf-

gereinigt. Dabei wurde anstelle von Bakterien der Ligationsansatz eingesetzt. Die Elution der DNA erfolgte abweichend vom Herstellerprotokoll in 30 µl (anstelle von 50 µl) Wasser. Die so aufgereinigte DNA wurde in einen Coating-Ansatz wie unter Beispiel 4
5 beschrieben eingesetzt (es wurden dann entsprechend weniger als die oben angegebenen 55 µl Wasser eingesetzt).

Die DNA wurde sodann wie unter Beispiel 4 beschrieben auf Tabak (*Nicotiana tabacum* cv. Petit Havanna CB199NTH X CB193NTH) Blätter
10 geschossen. Abweichend zu den Angaben in Beispiel 4 erfolgt die Selektion auf 30, 50 oder 100 mg/L Kanamycin.

Beispiel 12: Transformation von CB199NTHxCB193NTH Pflanzen mit pCB384-40

15

Zunächst wurde die Insertionssequenz in vitro aus dem Plasmid ausgeschnitten und diese anschließend zirkularisiert. 5 µg der DNA von pCB384-40 wurden dazu mit 10 U SalI (Roche) und 10U XhoI (Roche) in einem 20µl Restriktionsansatz unter den vom Hersteller
20 angegebenen Bedingungen für 2h inkubiert. Anschließend wurde der Ansatz 20 Minuten bei 65°C inkubiert. Über Nacht wurde bei 14°C eine Ligationsreaktion durchgeführt (Zugabe von 5µl 10x Ligasepuffer, 2U T4-Ligase (Roche) zum Restriktionsansatz und auffüllen auf 50 µl mit Wasser). Unter den gewählten Bedingungen wurde
25 überwiegend eine Rezirkularisierung der Fragmente beobachtet (nicht gezeigt). Die DNA wurde aus dem Ligationsansatz mit Hilfe des Wizard Plus Minipreps DNA Purification System (Promega) aufgereinigt. Dabei wurde anstelle von Bakterien der Ligationsansatz eingesetzt. Die Elution der DNA erfolgte abweichend vom Herstellerprotokoll in 30 µl (anstelle von 50 µl) Wasser. Die so
30 aufgereinigte DNA wurde in einen Coating-Ansatz wie unter Beispiel 4 beschrieben eingesetzt (es wurden dann entsprechend weniger als die oben angegebenen 55 µl Wasser eingesetzt). Die DNA wurde sodann wie unter Beispiel 4 beschrieben auf Tabak
35 (*Nicotiana tabacum* cv. Petit Havanna CB199NTH X CB193NTH) Blätter geschossen. Abweichend zu den Angaben in Beispiel 4 erfolgt die Selektion auf 30, 50 oder 100 mg/L Kanamycin.

Beispiel 13: Erstellen eines weiteren Vektors (pCB456-2) für die
40 Einbringung einer nicht natürlicherweise vorkommenden Erkennungsregion für die Homing Endonuklease I-PpoI und RE-Sequenzen (attB) für die Rekombinasen ϕ C31 und TP901-1 in das Plastom von Tabak

45 Ziel dieses Ansatzes war es (analog zu Beispiel 3) einen weiteren Vektor für die Plastidentransformation zu erstellen, welcher keine umfassenden Homologien zu Sequenzen im Plastidengenom

- besitzt. Der Selektionsmarker *aadA* ist in diesem Plasmid unter Kontrolle eines synthetischen Promotors, der abgeleitet ist von der Konsensus Sequenz für *E. coli* $\sigma 70$ Promotoren. Als 3'Ende wurde ein Bereich stromabwärts des 3'*psbA*-1 Gens aus *Synechocystis* genutzt. Im Unterschied zu dem bereits beschriebenen Vektor pCB199-3 wurde hier die RE-Sequenz direkt stromabwärts des *aadA* Gens, aber stromaufwärts der 3'*psbA*-1 Sequenz aus *Synechocystis* in das Molekül eingeführt. Damit kann man nach Insertion einer Insertionssequenz mit Hilfe einer Rekombinase ein Operon erstellen. Die Gene auf der Insertionssequenz können dann - optional - ohne Promotor auf der Insertionssequenz inseriert werden. Nach der Insertion gelangen entsprechende Gene der Insertionssequenz dann auch unter Kontrolle des synthetischen Promotors stromaufwärts des *aadA* Gens in der Masterpflanze. Dadurch kann man - optional - eine Operon Struktur bestehend aus dem *aadA* und den nachfolgend eingeführten Genen im Plastom erzeugen.

- Zur Herstellung des Plasmides pCB456-2 wurden sukzessive verschiedene Elemente in den Basisvektor pCB42-94 kloniert:

- a) Synthetischer Promotor (komplementär bp 1226-1260)
- b) Ribosomenbindestelle (komplementär bp 1214-1218)
- c) *aadA* Gen (komplementär bp 414-1208)
- 25 d) Core-Erkennungsregion für die Homing Endonuklease I-PpoI (komplementär bp 331-345)
- e) *attB* RE-Sequenz für die Rekombinasen ϕ C31 (bp 285-324) und TP901-1 (komplementär bp 276-176)
- f) 3'*psbA*-1 aus *Synechocystis* (komplementär bp 19-155)

- 30 Der so erhaltene Vektor vermittelt auch in *E. coli* eine Spectinomycin Resistenz. Anstelle der multiplen Klonierungsstelle im Basisvektor pCB42-94 (SEQ ID NO: 1) für die Plastidentransformation enthält dieser Vektor mit der Bezeichnung pCB456-2 die aufgeführten Elemente im Rahmen der Nukleinsäuresequenz mit der SEQ ID NO: 50. Auch hier ist die gesamte MCS ersetzende Sequenz (von *SacI* bis *KpnI*) angegeben.

- 40 Beispiel 14: Erstellen von überwiegend homotransplastomen Masterpflanzen, die eine nicht natürliche DSB-Erkennungssequenz und nicht natürliche RE-Sequenzen (*attB*) enthalten

- Analog zu pCB199-3 in Beispiel 4 wurde der Vektor pCB456-2 in die Plastiden von Tabak eingebracht. Abweichend von der Beschreibung in Beispiel 4 wurden die erhaltenen Sprosse jedoch auf Anzuchtmedium, welches 30 g/L Saccharose enthielt (anstelle der in Bei-

101

spiel 4 angegebenen 10 g/L), kultiviert. Die resultierenden Pflanzen wurden mit CB456NTH bezeichnet. Mittels Southern-Hybridisierung wurden von den Spectinomycin resistenten Pflanzen, die nach der Transformation erhalten wurden, 2 Linien (CB456NTH-1 und -15) identifiziert, die die Insertionssequenz aus pCB456-2 in ihr Plastom eingebaut haben. In dem Southern-Experiment wurde eine Sonde genutzt (SEQ ID NO: 58), welche gegen ein Fragment der 16SrDNA gerichtet war. Diese Sonde war geeignet, um aus mit EcoRI verdauter DNA, die dem Wildtyp entspricht, ein Fragment von etwa 3,1 kb nachzuweisen. Im Gegensatz dazu wurde ein etwa 1,7 kb großes Fragment detektiert, wenn in die entsprechenden Plastomkopien die Insertionssequenz aus pCB456-2 eingebaut worden war (siehe Fig. 11).

15 Beispiel 15: Erstellen einer Sequenz der TP901-1 Rekombinase mit modifizierter Aminosäuresequenz

Mit pCB127-2 als Matrize und den Oligonukleotiden p354 (SEQ ID NO: 51) sowie p355 (SEQ ID NO: 52) als Primer wurde eine PCR-Reaktion durchgeführt, um eine Ribosomenbindestelle stromaufwärts der für die TP901-1 Rekombinase kodierenden Sequenz einzubauen und die beschriebene Abweichung in der Aminosäuresequenz (siehe Beispiel 5.2) zu korrigieren. Der amplifizierte 5' Bereich der Rekombinase wurde in pCR4Blunt TOPO (Invitrogen) kloniert und anschließend als SpeI/HindIII Fragment in den mit XbaI und HindIII behandelten Vektor pCB127-2 ligiert. Der resultierende Vektor wurde pCB487-20 genannt. Das Insert in diesem Vektor wird durch SEQ ID NO: 60 beschrieben. Er umfasst eine Ribosomenbindestelle und kodiert für die TP901-1 Rekombinase (mit nativer Aminosäuresequenz; SEQ ID NO: 61).

p354: 5'-CGGATCCGGAGGCATCATGACTAAGAAAGTAGCAATCTATACACG-3'
(SEQ ID NO: 51)

35 p355: 5'-ATAAGCTTTATTCTCGATATCG-3' (SEQ ID NO: 52)

40 Beispiel 16: Erstellen eines Vektors, der für die Integration mit Hilfe von Rekombinasen in das Plastom von Masterpflanzen CB456NTH geeignet ist und zusätzlich einen Schnitt durch ein DSBI-Enzym ermöglicht

Sukzessive wurden die verschiedenen Elemente, die für einen solchen Vektor notwendig sind, in das Rückgrat des pBLUESCRIPT kloniert. Die angegebene Sequenz ersetzt die multiple Klonierungsstelle des pBLUESCRIPT von KpnI bis SacI. Das Rückgrat des Vektors entspricht dem des pBLUESCRIPT. Der erhaltene Vektor

wurde mit pCB557-1 bezeichnet. Das Insert wird durch
SEQ ID No: 53 beschrieben und umfasst nachfolgende Elemente:

- 5 a) Position Basenpaar 150-460: homologe Region zu Plastom-sequenzen der Masterpflanze CB456NTH - nach der Integration der hier betrachteten Insertionssequenz mittels Φ C31 oder TP901 Rekombinase liegt dieser Bereich dupliziert als direkter „repeat“ (gleiche Orientierung) zu den homologen Sequenzen aus CB456NTH vor. Dies erlaubt die intrachromosomale Rekombination nach der Insertion der Insertionssequenz, was zur Eliminierung der I-PpoI Erkennungsregion aus den bereits durch Insertion modifizierten Plastomkopien der Masterpflanze führt. In dieser Region homolog zu Sequenzen in CB456NTH enthalten sind von Basenpaar 150-244 (komplementär) 10 Teile des offenen Leserahmens 131 (ORF131) sowie von Basenpaar 337-431 das trnV Gen aus dem Tabakplastom der Masterpflanzen. 15
- b) attP Region der TP901 Rekombinase (bp 7-85)
- c) attP Region der Φ C31 Rekombinase (komplementär bp 96-145)
- 20 d) Synthetischer Promotor (bp 462-495)
- e) Ribosomenbindestelle (bp 514-518)
- f) nptII Gen (bp 523-1326)
- g) Ribosomenbindestelle (bp 1333-1337)
- h) I-PpoI codierendes Gen (bp 1342-1833)
- 25 Beispiel 17 Erstellen eines Vektors, der für die Integration mit Hilfe von Rekombinasen in das Plastom von Masterpflanzen CB456NTH geeignet ist und keinen zusätzlichen Schnitt durch ein DSBI-Enzym ermöglicht
- 30 Zum Vergleich zu pCB557-1 wurde ein Vektor erstellt, dessen Insertionssequenz für kein funktionelles DSBI-Enzym mehr kodiert. Dieser Vektor wurde wie folgt erstellt. Vektor pCB557-1 wurde mit dem Restriktionsenzym EcoO109I behandelt. Dadurch wurde ein 230 bp Fragment aus dem Gen kodierend für die Homing-Endonuklease 35 I-PpoI entfernt. Der verbleibende Vektoranteil wurde in einer Ligationsreaktion wieder rezirkularisiert. Der resultierende Vektor wurde pCB559-6 genannt.

- 40 Beispiel 15 Erstellen eines Vektors zur transienten Expression der Φ C31 Rekombinase in Plastiden

Um die Φ C31-Rekombinase transient in Plastiden exprimieren zu können, wurde diese in dem Vektor pCB554-3 stromabwärts eines synthetischen Promotors kloniert. Im Vektor pCB554-3 sind die 45 folgenden funktionellen Elemente vorhanden.

103

- a) Synthetischer Promotor (bp 28-62)
- b) Ribosomenbindestelle (bp 70-74)
- c) Gen kodierend für die Φ C31 Rekombinase (bp 79-1920)

5 Das Insert in pCB554-3 ist durch SEQ ID NO: 54 beschrieben.
Das Vektorrückgrat entspricht dem des pBLUESCRIPT.

Beispiel 16: Erstellen eines *in vitro* PCR-Produktes zur transien-
ten, plastidären Expression der TP901-Rekombinase

10

Das Gen kodierend für die TP901-1 Rekombinase aus pCB487-20 wurde
in geeigneter Orientierung mit einem synthetischen Promotor
fusioniert. Anschließend wird das Fusionsprodukt mittels PCR
amplifiziert; das entstehende Produkt wird mit PCR TP-Rek

15 bezeichnet. Die in der PCR einzusetzenden Primer p268
(SEQ ID NO: 55) und p72 (SEQ ID NO: 56) haben die Sequenz:

p268: TCGACTTGACATTCACTCTTCAATTATCTATAATGATACATG (SEQ ID NO: 55)

p72: CCTCGAGCTGCAGTTAAGCGAGTTGG (SEQ ID NO: 56)

20

Das resultierende PCR Produkt (TP-Rek) umfasst 1527 bp
(SEQ ID NO: 57) und umfasst folgende Elemente:

- a) Synthetischer Promotor (bp 6-40)
- 25 b) Ribosomenbindestelle (bp 48-52)
- c) Gen kodierend für die TP901 Rekombinase (bp 57-1514)

Beispiel 17: Transformationen von CB456NTH Masterpflanzen
zur Integration einer Insertionssequenz mittels
Erkennungsstellen-spezifischer Rekombination

30

In diesen Versuchen soll die Insertionssequenz oder das gesamte
Insertionskonstrukt durch die transiente Expression der Rekombi-
nase in Plastiden in die Plastomsequenz der Masterpflanzen

- 35 CB456NTH eingebaut werden. Das Aufbringen der entsprechenden
DNA auf die Goldpartikel erfolgt analog zu Beispiel 4. Folgende
Kombinationen von DNA werden jeweils gleichzeitig in die
Plastiden der Masterpflanzen CB456NTH eingebracht. Wird nur
die Insertionssequenz in die Plastiden eingebracht, wird diese
40 zunächst wie in Beispiel 12 beschrieben *in vitro* aus dem Inser-
tionskonstrukt heraus geschnitten (im Fall von pCB559-6 und
pCB557-1 erfolgt dies mit Hilfe der Restriktionsendonuklease
KpnI) und zirkularisiert. Wird das gesamte, unbehandelte Inser-
tionskonstrukt in der Transformation eingesetzt, wird auch das
45 Vektorrückgrat in die Plastiden eingebaut.

	Konstrukt zur transiente Expression der Rekombinase	Insertions- sequenz	Insertions- konstrukt
	pCB554-3	pCB557-1	-
	pCB554-3	-	pCB557-1
5	pCB554-3	pCB559-6	-
	pCB554-3	-	pCB559-6
	PCR TP-Rek	pCB557-1	-
	PCR TP-Rek	-	pCB557-1
10	PCR TP-Rek	pCB559-6	-
	PCR TP-Rek	-	pCB559-6

Der Transformationsprozess mittels Partikelkanone erfolgt
 15 im wesentlichen wie in Beispiel 14 für die Erstellung von
 CB456NTH beschrieben. Als wesentliche Unterschiede zu der
 obigen Beschreibung wird das Beschussmedium durch Regenerations-
 medium ersetzt und anstelle von Spectomycin wird Kanamycin als
 selektives Agens genutzt. In Kürze: Die beschossenen Explantate
 20 verbleiben für 2 Tage auf Regenerationsmedium und werden
 anschließend in etwa 0,5 cm² große Segmente zerschnitten. Diese
 werden für 4 bis 6 Wochen auf Regenerationsmedium zuzüglich
 30mg/L Kanamycin inkubiert. Anschließend werden die Explantate
 für 7 bis 10 Wochen auf Regenerationsmedium zuzüglich 50 mg/L
 25 Kanamycin inkubiert. Danach werden die Explantate - optional -
 auf Regenerationsmedium zuzüglich 80mg/L Kanamycin inkubiert.
 Entstehende Pflanzen werden wiederum auf Anzuchtmedium (zuzüg-
 lich Kanamycin) gesetzt bis sie Wurzeln bilden. Die Bewurzelung
 kann ggf. auch in Abwesenheit des selektiven Agens erfolgen.
 30

35

40

45

Patentansprüche

1. Verfahren zur sequenzspezifischen Integration einer DNA-
5 Sequenz in die plastidäre DNA eines pflanzlichen Organismus
oder von diesem abgeleitete Zellen, dadurch gekennzeichnet,
dass
- a) die plastidären DNA-Moleküle des besagten pflanzlichen
10 Organismus oder von diesem abgeleitete Zelle mindestens
eine Rekombinaseerkennungssequenz R1 enthalten und
- b) i) mindestens ein Transformationskonstrukt umfassend
eine Insertionssequenz zumindest einseitig flankiert
15 von mindestens einer weiteren Rekombinationssequenz
R2 sowie
- ii) mindestens eine sequenzspezifische Rekombinase
geeignet zur Induktion von Rekombinationsereignissen
20 an besagten Rekombinationssequenzen R1 und R2
- in mindestens einem Plastid besagten pflanzlichen
Organismus oder von diesem abgeleiteten Zellen zusammen-
gebracht werden, und
- 25 c) die Insertionssequenz in die plastidäre DNA unter
Rekombination von R1 mit R2 insertiert, und
- d) Pflanzen oder Zellen isoliert werden, bei denen die
30 Insertionssequenz in die plastidären DNA-Moleküle
insertiert wurde.
2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, dass die
35 Funktionalität der durch Rekombination von R1 mit R2 ent-
stehenden Rekombinationssequenzen so vermindert wird, dass
die Insertion der Insertionssequenz gegenüber der Excision
derselben um mindestens den Faktor zehn bevorzugt ist.

40

45

Zeichn.

106

3. Verfahren nach einem der Ansprüche 1 oder 2, dadurch gekennzeichnet, dass das Transformationskonstrukt mindestens eines der nachfolgenden Elemente beinhaltet
- 5 i) Expressionskassette für eine Rekombinase
- ii) Positiven Selektionsmarkern
- iii) Negativen Selektionsmarkern
- 10 iv) Reportergen
- v) Replikationsursprüngen
- 15 vi) Multiple Klonierungsregionen
- vii) "Border"-Sequenzen für Agrobakterium-Transfektion
- viii) Sequenzen, die eine homologe Rekombination bzw.
- 20 Insertion in das Genom eines Wirtsorganismus ermöglichen
- ix) Expressionskassette für ein Enzym, geeignet zur
- 25 Induktion von DNA-Doppelstrangbrüchen an der
- Erkennungssequenz zur gezielten Induktion von
- DNA-Doppelstrangbrüchen
4. Verfahren nach einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, dass die Rekombinase ausgewählt ist aus der Gruppe
- 30 der Rekombinasen bestehend aus Cre, Flp, Φ C31 Rekombinase, TP901-1 Rekombinase, Gin, Cin, Pin, λ -Rekombinase, R4 Integrase und Hin Invertase.
5. Verfahren nach einem der Ansprüche 1 bis 4, dadurch gekennzeichnet, dass die Rekombinase als Fusionsprotein mit einer
- 35 Plastidenlokalisationssequenz exprimiert wird.
6. Verfahren nach einem der Ansprüche 1 bis 5, dadurch gekennzeichnet, dass die Rekombinase kodiert ist durch eine Sequenz
- 40 gemäß SEQ ID NO: 41, 42, 46 oder 47.
7. Expressionskassette enthaltend eine Nukleinsäuresequenz kodierend für eine Rekombinase ausgewählt aus der Gruppe der
- 45 Resolvasen/Invertasen unter Kontrolle eines in pflanzlichen Plastiden aktiven Promotors.

107

8. Expressionskassette enthaltend eine Nukleinsäuresequenz kodierend für Fusionsprotein aus einer Plastiden-lokalisationssequenz und einer Rekombinase ausgewählt aus der Gruppe der Resolvasen/Invertasen unter Kontrolle eines im pflanzlichen Zellkern aktiven Promotors.
9. Expressionskassette nach einem der Ansprüche 7 oder 8, dadurch gekennzeichnet, dass die Rekombinase ausgewählt ist aus der Gruppe bestehend aus Φ C31 Rekombinase, R4 Rekombinase, Hin Rekombinase, TP901-1 Rekombinase, Gin, Cin und Pin.
10. Pflanzlicher Organismus erhalten nach einem Verfahren gemäß einem der Ansprüche 1 bis 6.
11. Pflanzlicher Organismus enthaltend eine Expressionskassette gemäß einem der Ansprüche 7 bis 9.
12. Pflanzlicher Organismus nach einem der Ansprüche 10 oder 11 ausgewählt aus der Gruppe bestehend aus *Arabidopsis thaliana*, Tabak, *Tagetes*, Weizen, Roggen, Gerste, Hafer, Raps, Mais, Kartoffel, Zuckerrübe, Soja, Sonnenblume, Kürbis oder Erdnuss.
13. Zellkulturen, Organe, Gewebe, Teile oder transgenes Vermehrungsgut abgeleitet von einem pflanzlichen Organismus nach den Ansprüchen 10 bis 12.
14. Verwendung eines Organismus nach einem der Ansprüche 10 bis 12 oder von diesem abgeleitete Zellkulturen, Organe, Gewebe, Teile oder transgenes Vermehrungsgut nach Anspruch 13 als Nahrungs-, Futtermittel oder Saatgut oder zur Herstellung von Pharmazeutika oder Feinchemikalien.

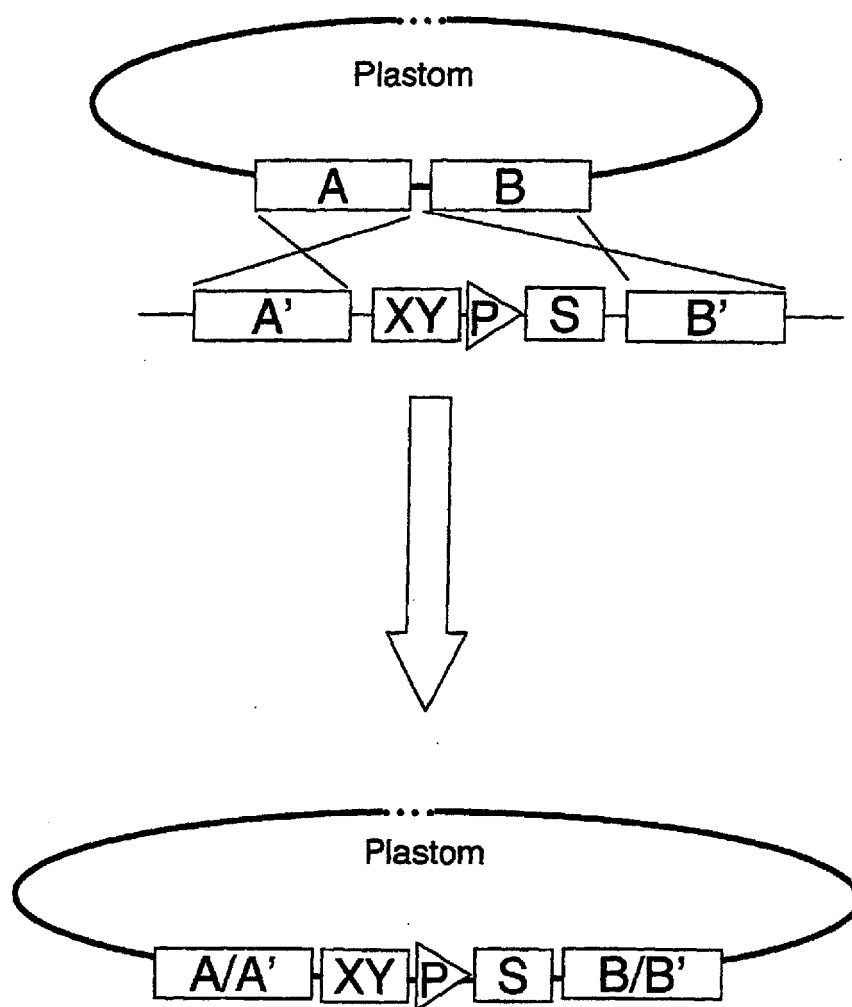


Fig. 1A

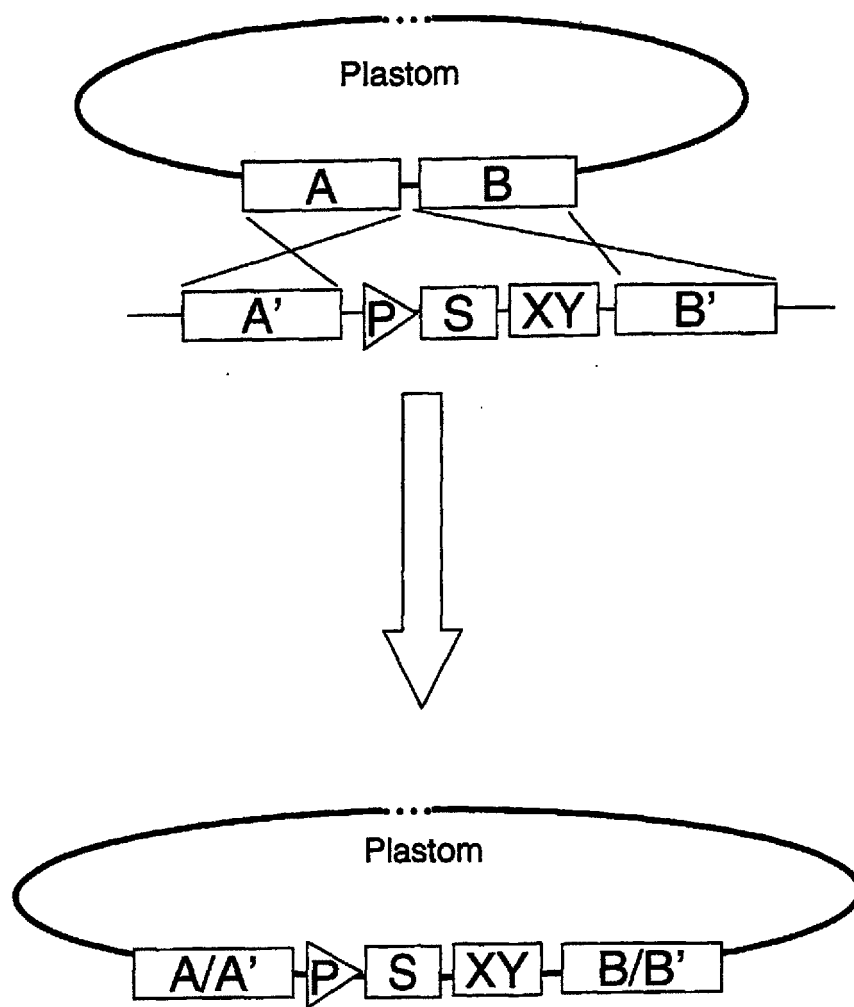


Fig. 1B

3/14

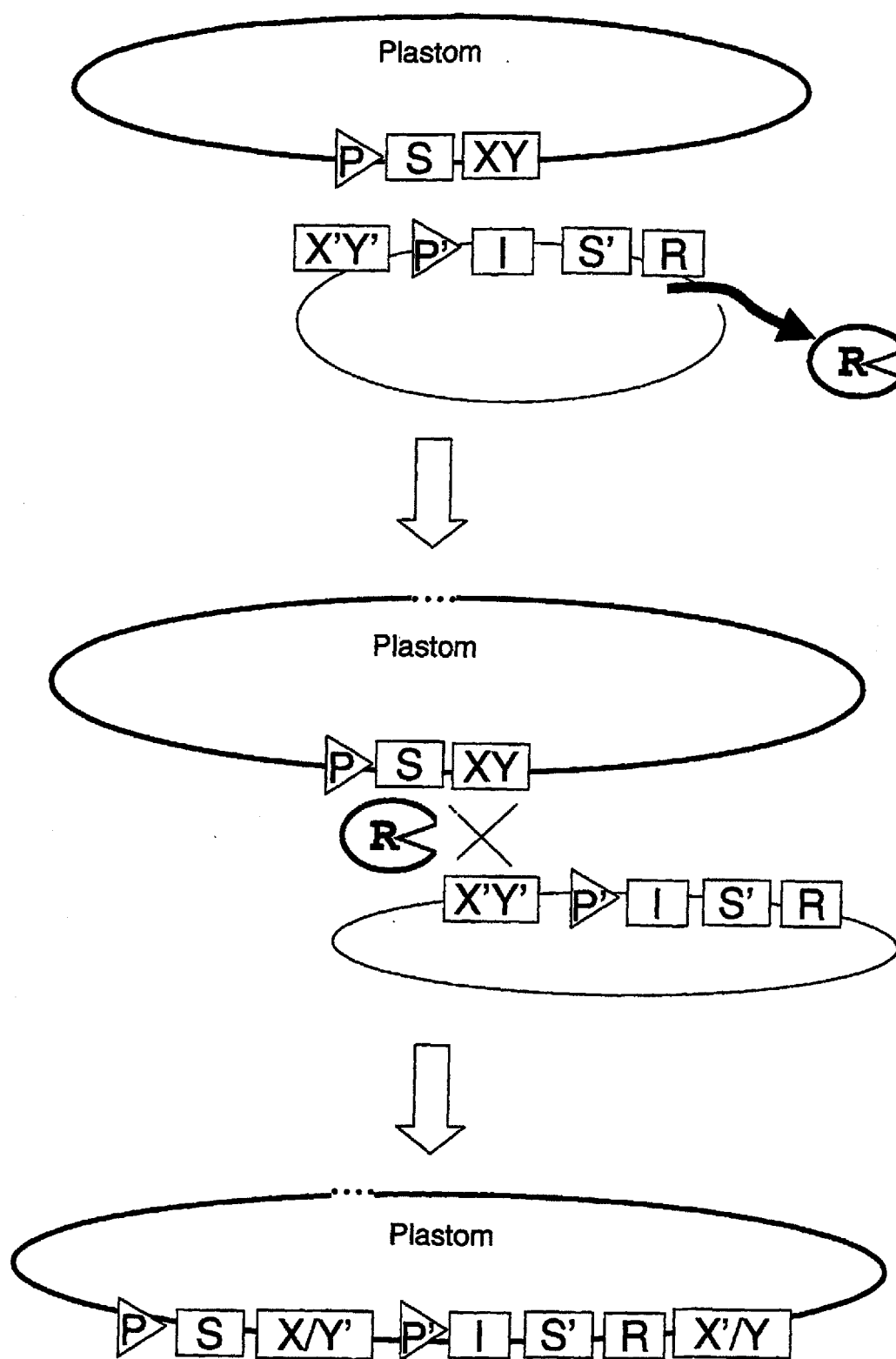


Fig. 2

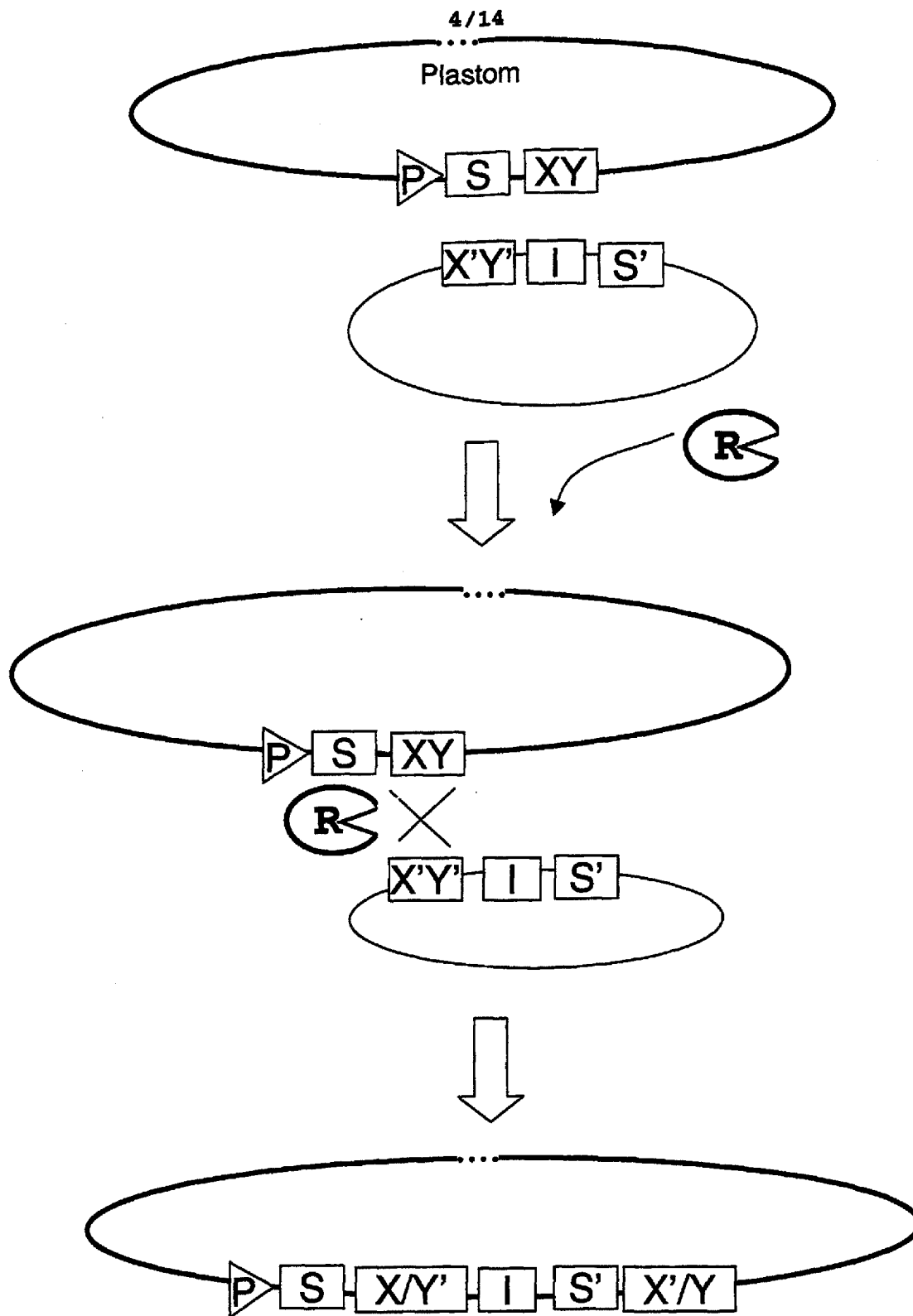


Fig. 3

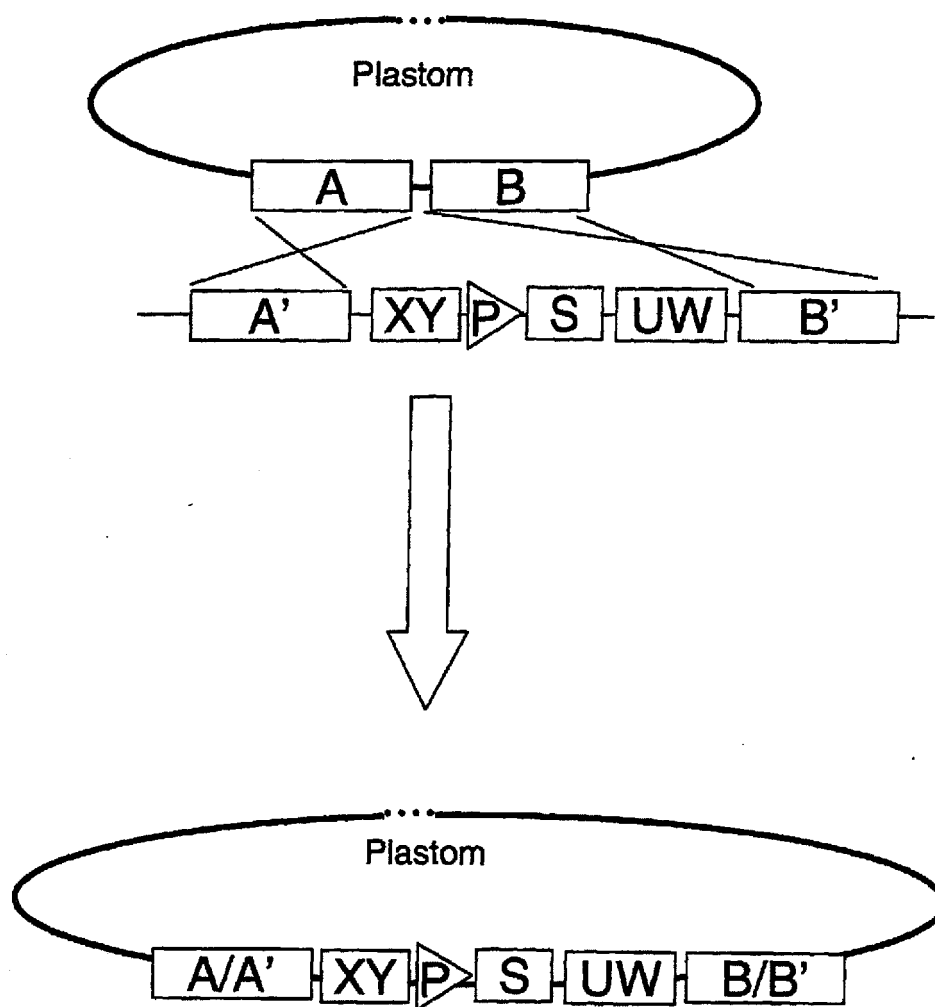


Fig. 4

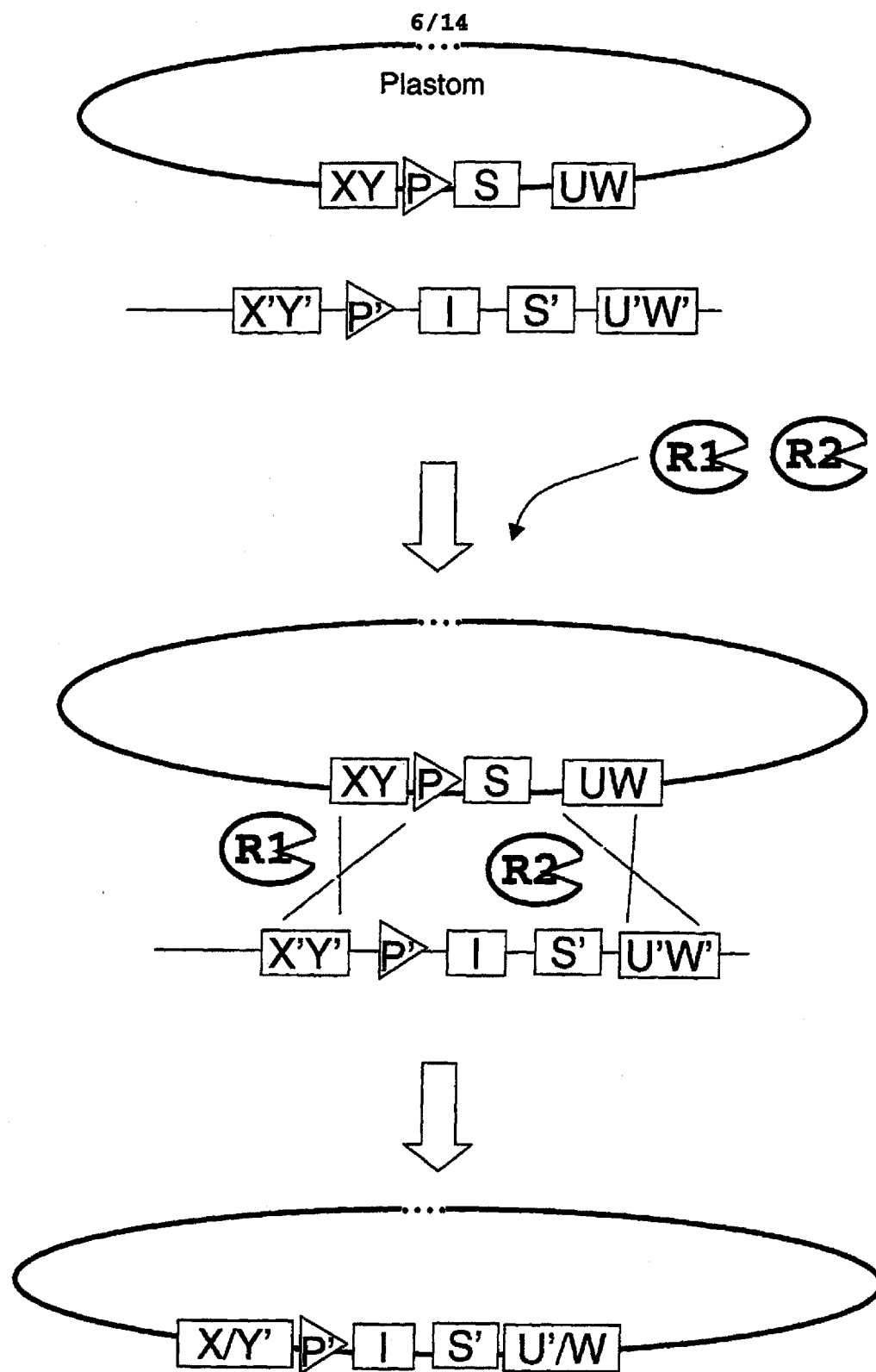


Fig. 5

7/14

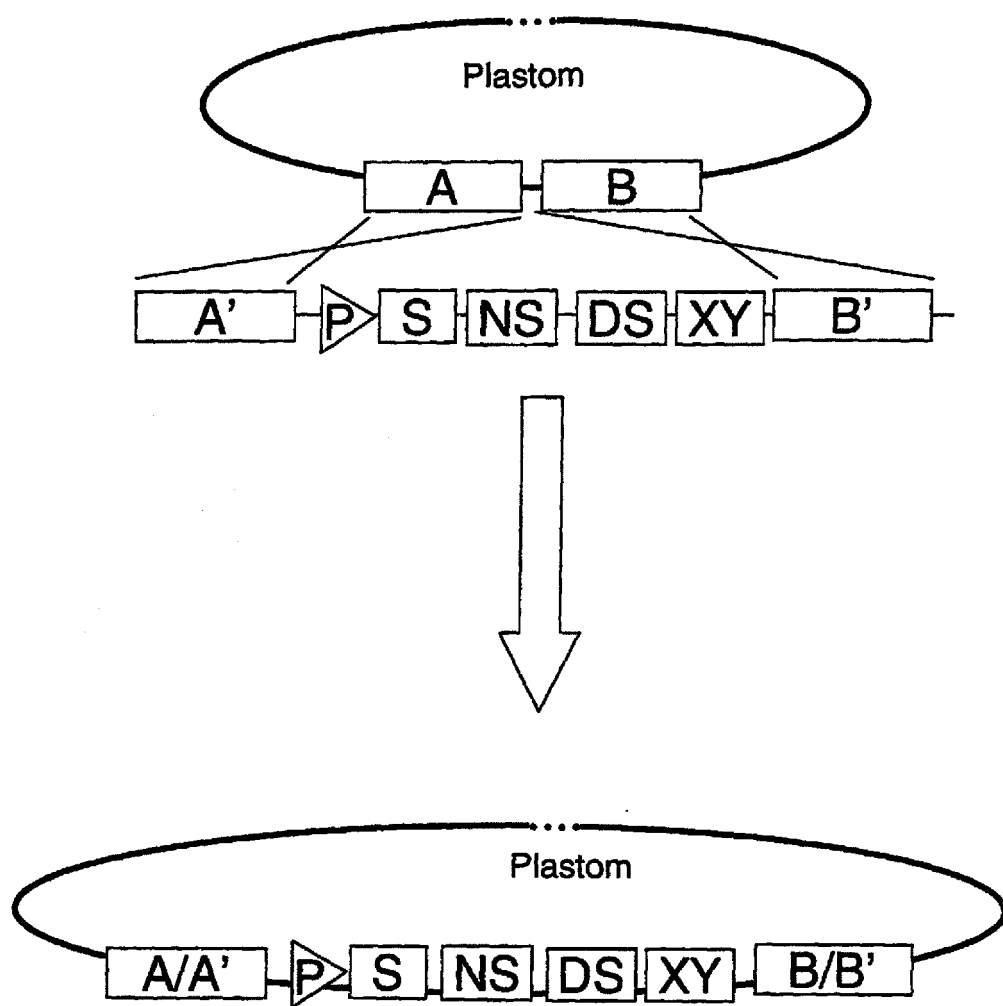


Fig. 6

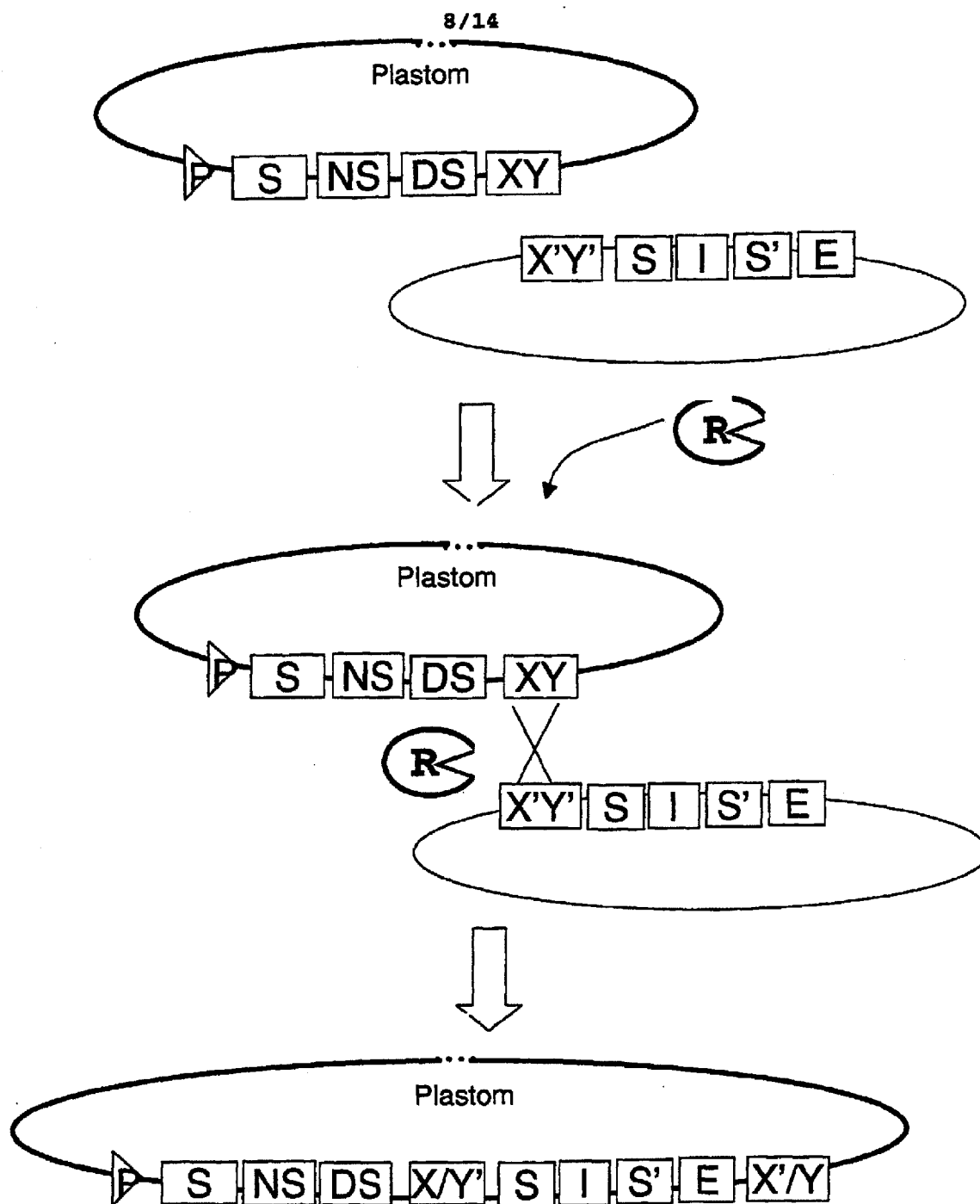


Fig. 7A

9/14

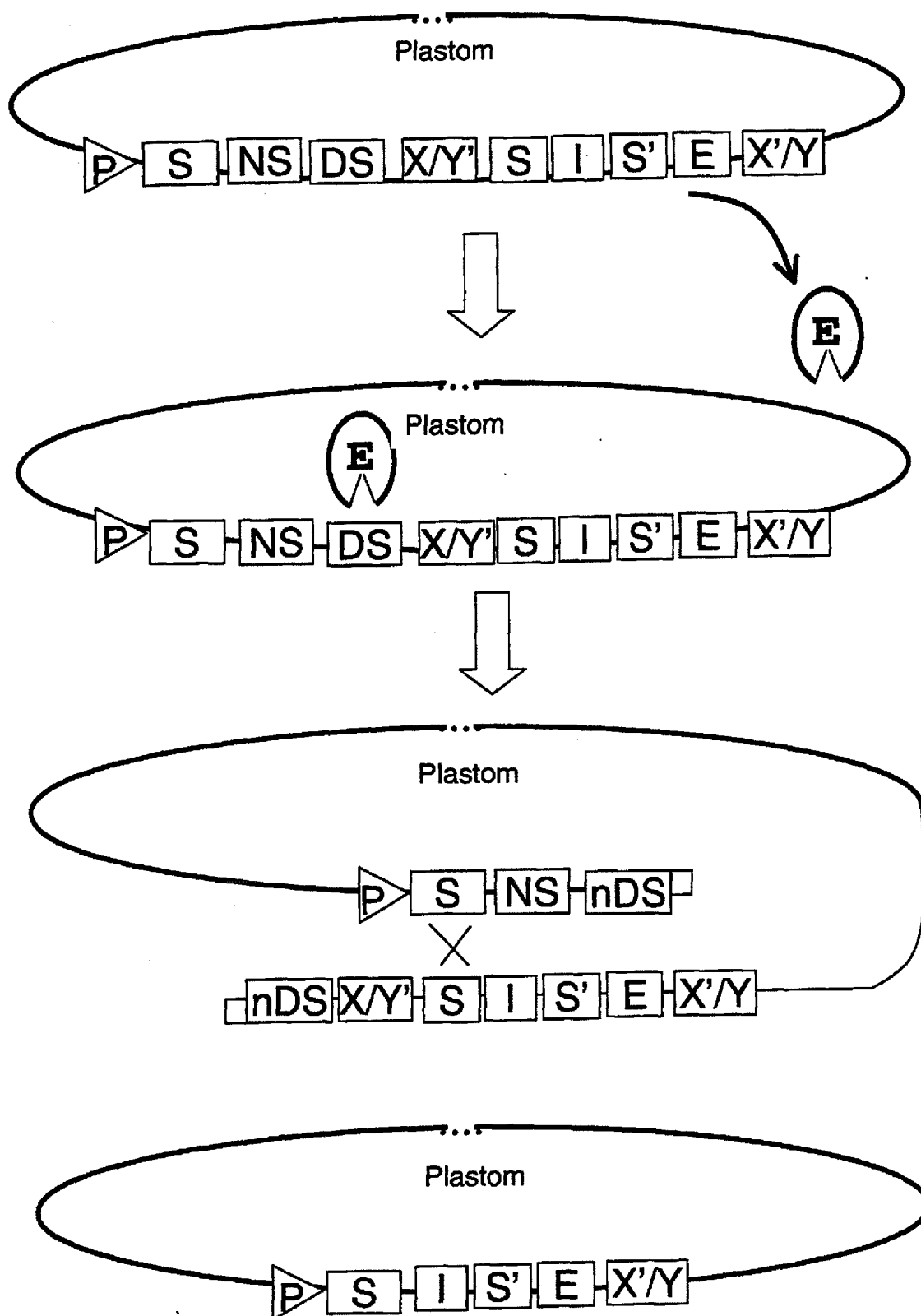


Fig. 7B

10/14

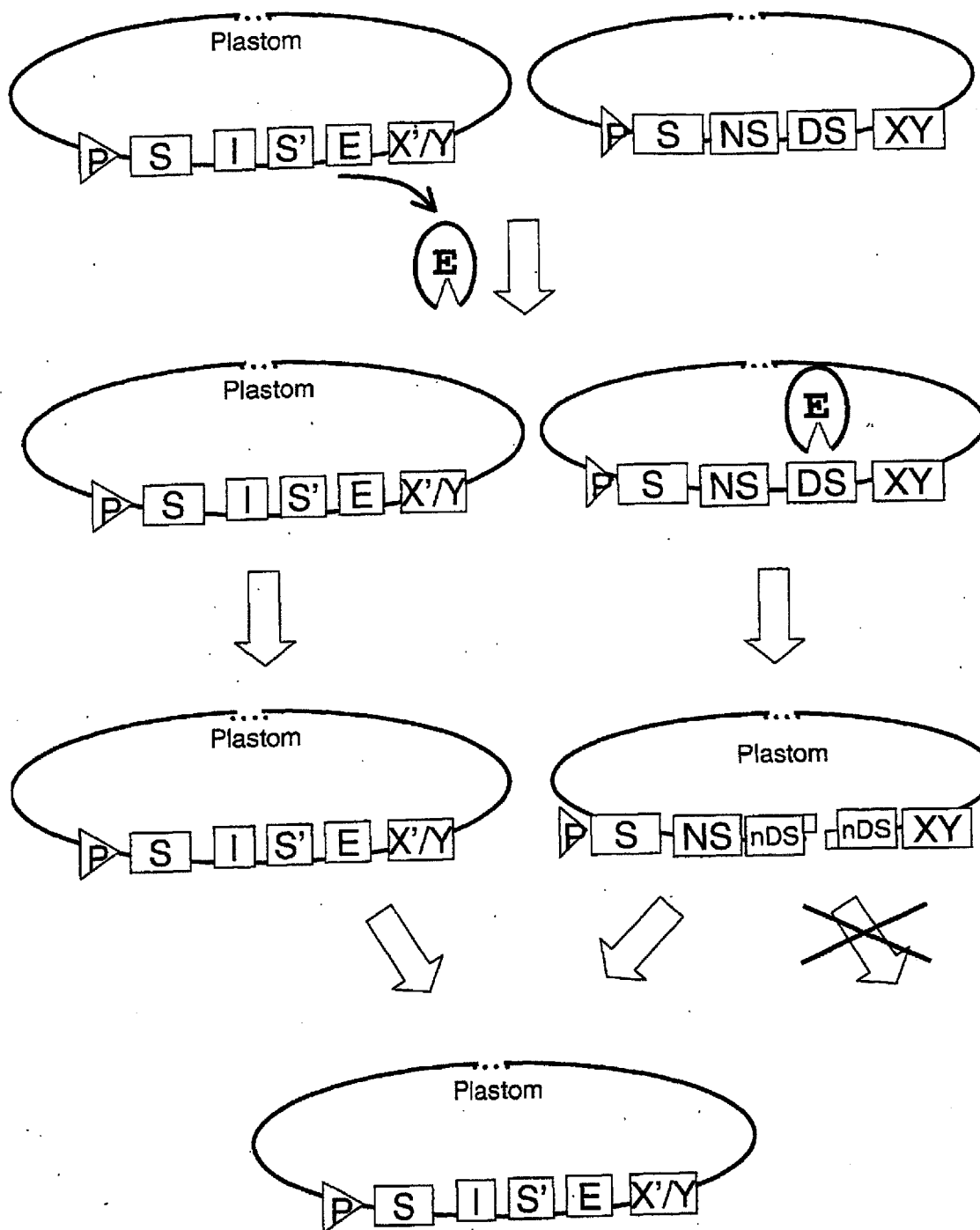


Fig. 7C

11/14

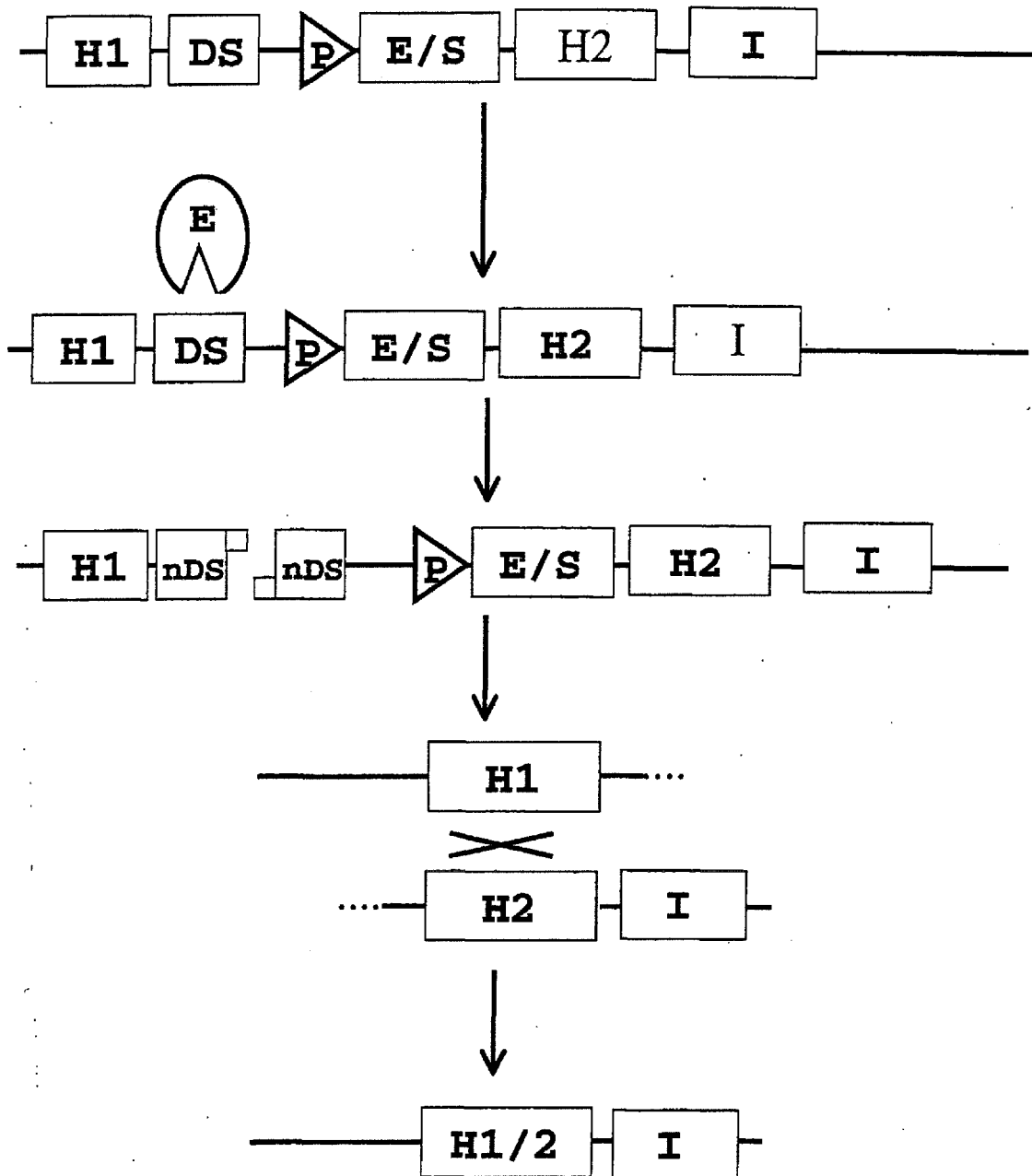


Fig. 8

12/14

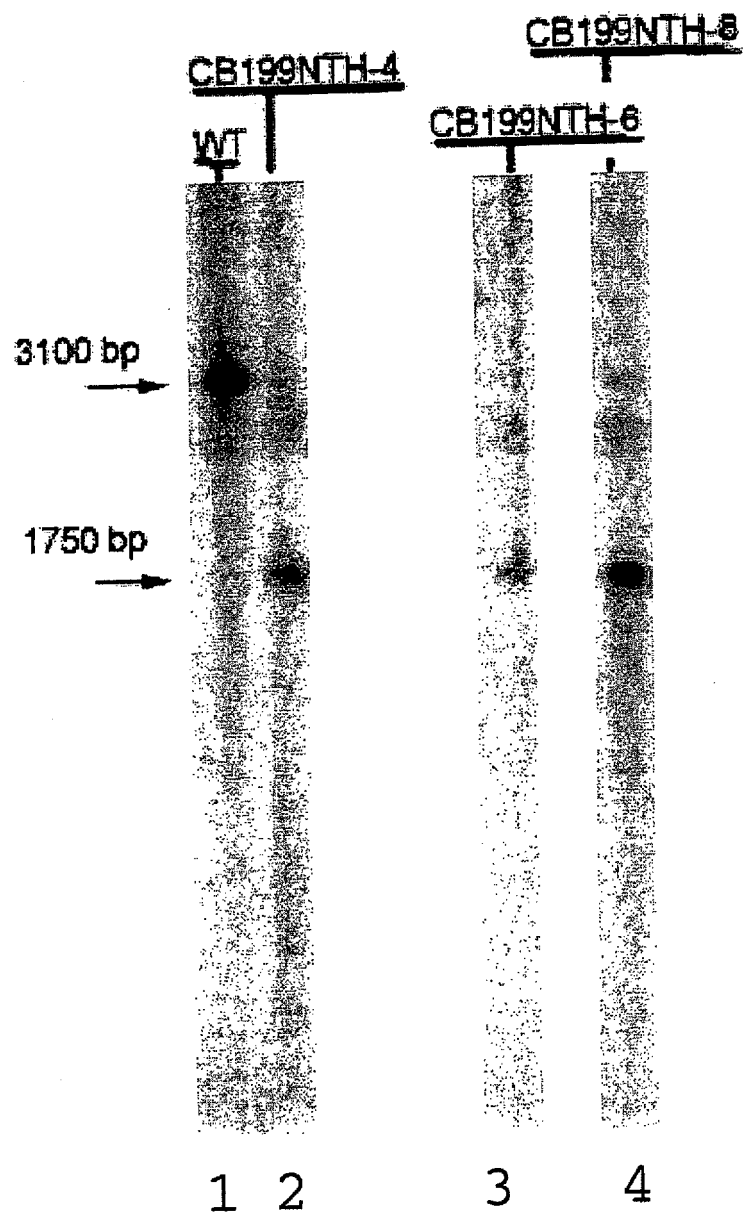
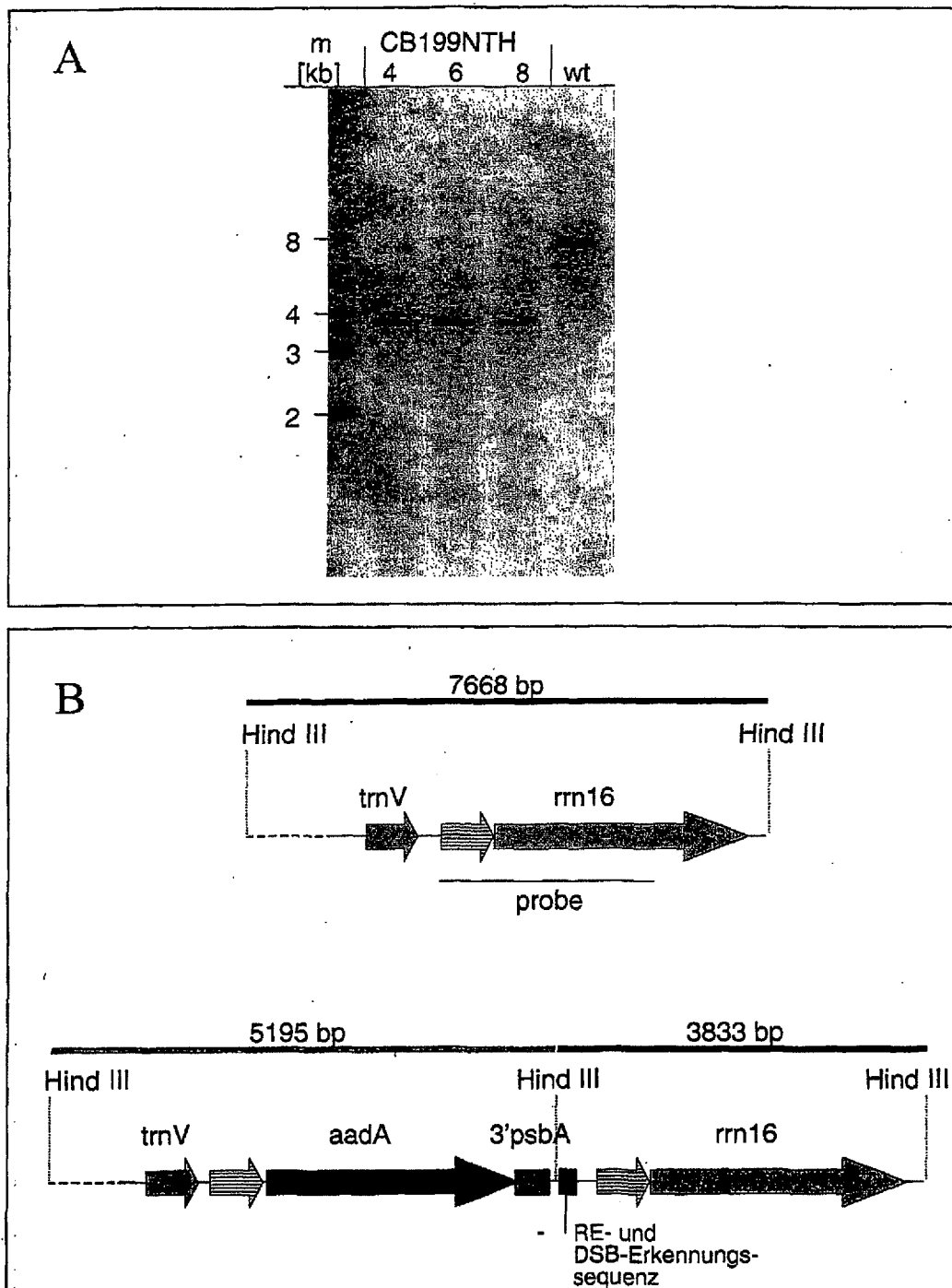
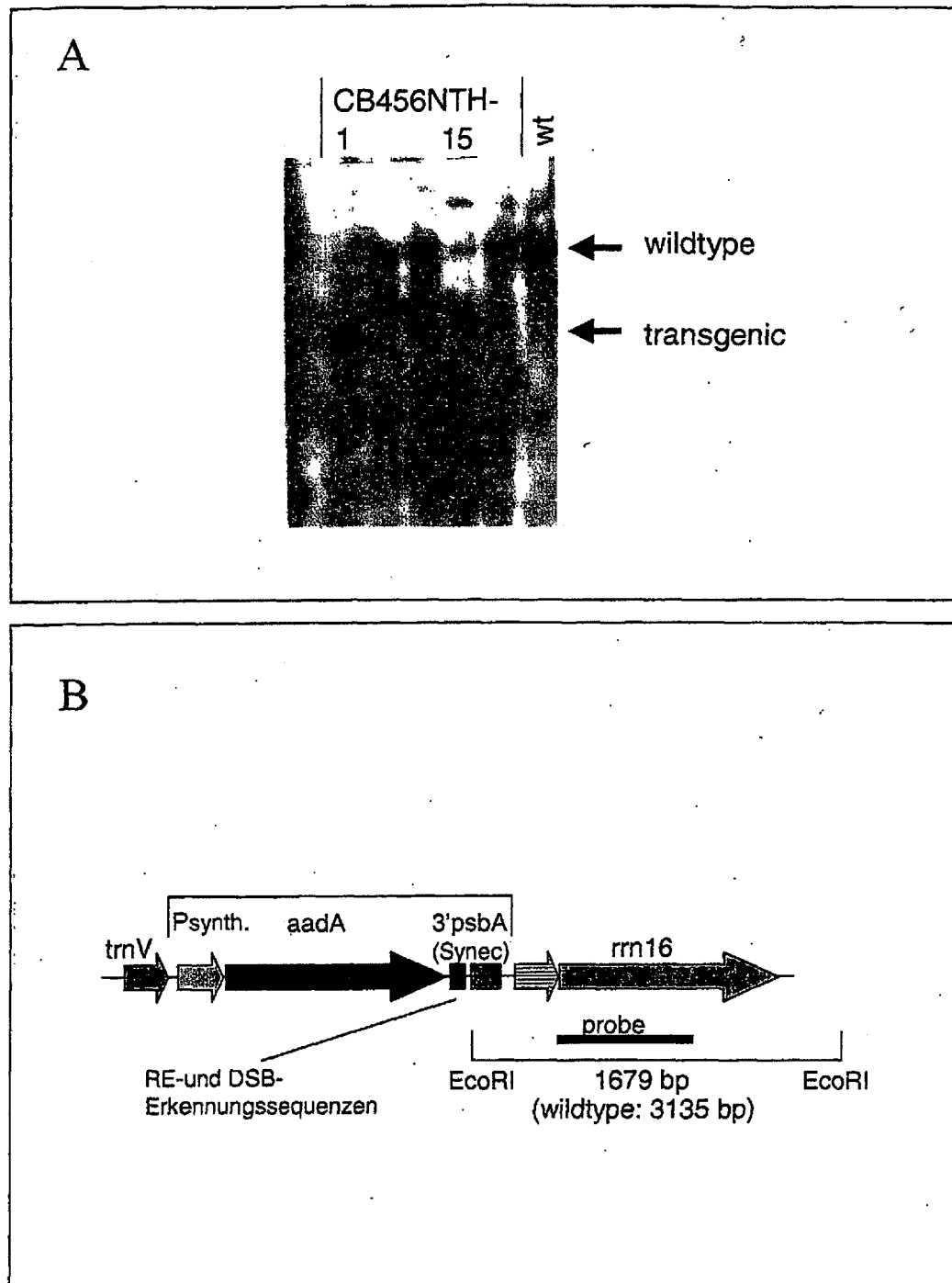


Fig. 9

**Fig. 10**

**Fig. 11**

SEQUENZPROTOKOLL

<110> SunGene GmbH & Co.KGaa

<120> Verfahren zur Transformation von pflanzlichen Plastiden

<130> PF0817000029

<140>

<141>

<160> 61

<170> PatentIn Ver. 2.1

<210> 1

<211> 4363

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: plastid transformation plasmid pCB42-94

<220>

<221> misc_feature

<222> (55)..(1405)

<223> right targeting region

<220>

<221> mutation

<222> (346)

<223> mutation causing streptomycin resistance

<220>

<221> mutation

<222> (68)

<223> mutation causing spectinomycin resistance

<220>

<221> misc_feature

<222> (1404)..(1511)

<223> multiple cloning site

<220>

<221> misc_feature

<222> (2629)..(3417)

<223> Ampicilin resistance

<400> 1

```
gcgccaata cgcaaaccgc ctctccccgc gcgttgccgc attcattaat gcagtgatga 60
cttgacggca tcctcacctt cctccggctt atcaccggca gtctgttcag ggttcctaac 120
tcaacgatgg caactaaaca cgagggttgc gtcgttgcg ggacttaacc caacacctta 180
cggcacgagc tgacgacagc catgcaccac ctgtgtccgc gttcccgaag gcacccctct 240
ctttcaagag gattcgcggc atgtcaagcc ctggtaaggt tcttcgcttt gcacgaatt 300
aaaccacatg ctccaccgct tgtgcgggcc cccgtcaatt ccttttagtt tcattcttgc 360
gaacgtactc cccaggcggg atacttaacg cgtagctac agcactgcac gggtcgatac 420
gcacagcgcc tagtatccat cgtttacggc taggactact ggggtatcta atcccattcg 480
ctcccctagc tttcgtctct cagtgtcagt gtcggcccag cagagtgcct tcgccgttgg 540
tgttctttcc tagctctacg catttcaccg ctccaccgga aattccctct gccctaccg 600
tactccagct tggtagtttc caccgcctgt ccagggttga gccctgggat ttgacggcgg 660
acttaaaaag ccacctacag acgctttacg cccaatcatt ccggataacg cttgcacct 720
ctgtattacc ggggtgctg gcacagagtt agccgatgct tattccccag ataccgtcat 780
tgcttcttct cggggaag aagttcacga cccgtgggcc ttctacctcc acgcggcatt 840
gtccgctcag gctttcgccc attgcggaaa attcccact gctgcctccc gtaggagtct 900
```


gggcccgtgtc	tcagtcccag	tgtggctgat	catcctctcg	gaccagctac	tgatcatcgc	960
cttggttaagc	tattgcctca	ccaactagct	aatcagacgc	gagcccctcc	tcgggaggat	1020
tcctcctttt	gctcctcagc	ctacggggta	ttagcagccg	tttccagctg	ttgttcccc	1080
cccaagggca	ggttcttacg	cgttactcac	ccgtccgccca	ctggaaacac	cacttcccgt	1140
ccgacttgca	tgtgttaagc	atgccgccag	cgttcatcct	gagccaggat	cgaactctcc	1200
atgagattca	tagttgcatt	acttatagct	tccttggtcg	tagacaaagc	ggattcggaa	1260
ttgtctttca	ttccaaggca	taacttgat	ccatgcgctt	catattcgcc	cggagtctgc	1320
tcccagaaat	atagccatcc	ctgccccctc	acgtcaatcc	cacgagcctc	ttatccattc	1380
tcattgaacg	acggcggggg	agcgagctcc	accgcggtgg	cggccgctct	agaactagt	1440
gatcccccg	gctgcaggaa	ttcgatatca	agcttatcga	taccgtcgac	ctcgagggg	1500
ggcccgggtac	caaatccaac	tagaaaaact	cacattgggc	ttagggataa	tcaggctcga	1560
actgatgact	tccaccacgt	caaggtgaca	ctctaccgct	gagttatata	ccttccccgc	1620
cccatcgaga	aatagaactg	actaatccta	agtcaaagg	tcgagaaact	caacgccact	1680
attcttgaac	aacttggagc	cgggccttct	tttcgcacta	ttacggatat	gaaaataatg	1740
gtcaaaatcg	gattcaattg	tcaactgccc	ctatcgga	taggattgac	taccgattcc	1800
gaaggaaactg	gagttacatc	tcttttccat	tcaagagttc	ttatgcgttt	ccacgcccct	1860
ttgagacccc	gaaaaatgga	caaattcctt	ttcttaggaa	cacatacaag	attcgctact	1920
acaaaaagga	taatggtaac	cctaccatta	actacttcat	ttatgaattt	catagtaata	1980
gaaatacatg	tcctaccgag	acagaatttg	gaacttgcta	tcctcttgcc	tagcaggcaa	2040
agattttacct	ccgtggaaag	gatgattcat	tcggatcgac	atgagagtcc	aactacattg	2100
ccagaatcca	tgttgatat	ttgaaagagg	ttgacctcct	tgttctctc	atggtacact	2160
cctcttcccg	ccgagcccc	tttctcctcg	gtccacagag	acaaaatgta	ggactggtag	2220
caacaattca	tcagactcac	taagtccgga	tcactaacta	atactaactt	aataataatg	2280
tctaataat	ctaataat	agaaaatact	aataataatg	aaaagaactg	tcttttctgt	2340
atactttccc	cggttccggt	gctaccgagg	gcctcgtgat	acgcctattt	ttatagggtta	2400
atgtcatgat	aataatgggt	tcttagacgt	caggtggcac	ttttcgggga	aatgtgcgcg	2460
gaaccctgat	ttgttttatt	ttctaaatac	attcaaatat	gtatccgctc	atgagacaat	2520
aaccctgata	aatgcttcaa	taatattgaa	aaaggaagag	tatgagtatt	caacatttcc	2580
gtgtcgccct	tattcccttt	tttgccgcat	tttgcccttc	tgtttttgct	caccagaaa	2640
cgctggtgaa	agtaaaagat	gctgaagatc	agttgggtgc	acgagtgggt	tacatcgaac	2700
tggatctcaa	cagcggtaag	atccttgaga	gttttcgccc	cgaagaacgt	tttccaatga	2760
tgagcacttt	taaagtcttg	ctatgtggcg	cggtattatc	ccgtattgac	gccgggcaag	2820
agcaactcgg	tcgccgcata	cactattctc	agaatgactt	ggttgagtac	tcaccagtca	2880
cagaaaagca	tcttacggat	ggcatgacag	taagagaatt	atgcagtgtc	gccataacca	2940
tgagtataaa	cactgcccgc	aacttacttc	tgacaacgat	cggaggaccg	aaggagctaa	3000
ccgctttttt	gcacaacatg	ggggatcatg	taactcgcct	tgatcgttgg	gaaccggagc	3060
tgaatgaagc	cataccaaac	gacgagcgtg	acaccacgat	gcctgtagca	atggcaacaa	3120
cgttgcgcaa	actattaact	ggcgaactac	ttactctagc	ttcccggcaa	caattaatag	3180
actggatgga	ggcggataaa	gttgacggac	cacttctgct	ctcgccctt	ccggtgggt	3240
ggtttattgc	tgataaatct	ggagccggtg	agcgtgggtc	tcgcggtatc	attgcagcac	3300
tggggccaga	tggttaagccc	tcccgatcgc	tagttatcta	cacgacgggg	agtcaggcaa	3360
ctatggatga	acgaaataga	cagatcgctg	agatagggtg	ctcactgatt	aagcattggt	3420
aactgtcaga	ccaagtttac	tcataataac	tttagattga	tttaaaactt	catttttaac	3480
ttaaaaggat	ctaggtgaag	atcctttttg	ataatctcat	gaccaaaatc	ccttaacgtg	3540
agttttcgtt	ccactgagcg	tcagaccccg	tagaaaagat	caaaggatct	tcttgagatc	3600
ctttttttct	gcgcgtaatc	tgctgcttgc	aaacaaaaaa	accaccgcta	ccagcgggtg	3660
tttggttgcc	ggatcaagag	ctaccaactc	tttttccgaa	ggtaactggc	ttcagcagag	3720
cgcatatacc	aaatactgtc	cttctagtgt	agccgtagtt	aggccaccac	ttcaagaact	3780
ctgtagcacc	gcctacatac	ctcgcctctg	taatcctggt	accagtggct	gctgccagt	3840
gcgataagtc	gtgtcttacc	gggttgga	caagacgata	gttaccggat	aaggcgcagc	3900
ggtcgggctg	aacggggggg	tcgtgcacac	agcccagctt	ggagcgaacg	acctacaccg	3960
aactgagata	cctacacggt	gagctatgag	aaagcgccac	gcttcccga	gggagaaagg	4020
cggacaggta	tccggtaagc	ggcagggtcg	gaacaggaga	gcgcacgagg	gagcttccag	4080
ggggaaacgc	ctggtatctt	tatagtctcg	tcgggtttcg	ccacctctga	cttgagcgtc	4140
gattttttgtg	atgctcgtca	ggggggcgga	gcctatggaa	aaacgccagc	aacggggcct	4200
ttttacgggt	cctggccttt	tgctggcctt	ttgctcacat	gttctttcct	gcgttatccc	4260

3

ctgattctgt ggataaccgt attaccgct ttgagtgage tgataaccgt cgccgcagcc 4320
 gaacgaccga ggcgcagcag tcaagtgcgc aggaagcggg aga 4363

<210> 2

<211> 1359

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: Insert of
 vector pCB199-3

<400> 2

gagctctgat cacggaagat agctttggca aaaaaagcaa aaagcattta ccttgattga 60
 gatgttaatt gtgttgcaa ttatcagtat tttatttttg ctttttgtgc caaatttgat 120
 actagagctt cgggtgccag ggcgtgccct tgggtcccc gggcgctac tcgacgctac 180
 ctttaagagag tcaagcttct atattaccct gttatcccta gcgtactcga gaaaaaaga 240
 aaggagcaat agcacctctt tgatagaaca agaaaatgat tattgctcct ttcttttcaa 300
 aacctcttat agactaggcc aggaattatc tgcagttatt tgccaactac cttagtgatc 360
 tcgcctttca cgtagtggac aaattcttcc aactgatctg cgcgcgaggc caagcgatct 420
 tcttcttgtc caagataagc ctgtctagct tcaagtatga cgggctgata ctgggcccgc 480
 aggcgctcca ttgccagtc ggcagcgaca tcttcggcg cgattttgcc gggtactgcg 540
 ctgtaccaa tgccgggacaa cgtaagcact acatttcgct catcgccagc ccagtcgggc 600
 ggcgagttcc atagcgttaa ggtttcattt agcgcctcaa atagatcctg ttcaggaacc 660
 ggatcaaaga gttcctccgc cgctggacct accaaggcaa cgctatgttc tcttgctttt 720
 gtcagcaaga tagccagatc aatgtcgatc gtggctggct cgaagatacc tgcaagaatg 780
 tcattgcgct gccattctcc aaattgcagt tcgcgcttag ctggataacg ccacggaatg 840
 atgtcgtcgt gcacaacaat ggtgacttct acagcgcgga gaatctcgct ctctccaggg 900
 gaagccgaag ttccaaaag gtcgttgatc aaagctcgcc gcgttggttc atcaagcctt 960
 acggtcaccg taaccagcaa atcaatatca ctgtgtggct tcaggccgac atccactgcg 1020
 gagccgtaca aatgtacggc cagcaacgct ggttcgagat ggcgctcgat gacgccaact 1080
 acctctgata gttgagttga tacttcggcg ataaccgctt cagagccat ggctgtttgt 1140
 ggtgtcatgg ctgtttgtgg tgtoatgaat ccctccctac aactagatcc tcgcccggag 1200
 ttcgctccca gaaatatagc catccctgcc ccctcacgtc aatcccacga gcctcttctc 1260
 cattctcatt gaacgacggc gtcgaggggg ggcccgttac gtcgacgaag ttcctattcc 1320
 gaagttccta ttctcaagaa agtataggaa ctctgtacc 1359

<210> 3

<211> 1070

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: insert of
 vector pCB401-20

<400> 3

gagctctgat cacggaagat agctttggca aaaaaagcaa aaagcattta ccttgattga 60
 gatgttaatt gtgttgcaa ttatcagtat tttatttttg ctttttgtgc caaatttgat 120
 actagagctt cgggtgccag ggcgtgccct tgggtcccc gggcgctac tcgacgctac 180
 ctttaagagag tcaagcttct atattaccct gttatcccta gcgtactcga gctgcagtta 240
 tttgccaaact accttagtga tctcgcttct cacgtagtgg acaaattctt ccaactgatc 300
 tgcgcgcgag gccaaagcat cttcttcttg tccaagataa gcctgtctag cttcaagtat 360
 gacgggctga tactgggccc gcaggcgctc cattgccagc tcggcagcga catccttcgg 420
 cgcgattttg ccggttactg cgctgtacca aatgcgggac aacgtaagca ctacatttcg 480
 ctcatcgcca gccagtcgg ggcgcgagtt ccatagcgtt aaggtttcat ttagcgcttc 540
 aaatagatcc tgttcaggaa ccggatcaaa gagttcctcc gccgctggac ctaccaaggc 600
 aacgctatgt tctcttgctt ttgtcagcaa gatagccaga tcaatgtcga tcgtggctgg 660
 ctgaagata cctgcaagaa tgctattgcy ctgccattct ccaaattgca gttcgcgctt 720
 agctggataa cgccacggaa tgatgtcgtc gtgcacaaca atggtgactt ctacagcgcg 780
 gagaatctcg ctctctccag ggaagccga agtttccaaa aggtcgttga tcaagctcg 840

ccgcgttggt tcatcaagcc ttacgggtcac cgtaaccagc aaatcaatat cactgtgtgg 900
 cttcaggccg ccatccactg cggagccgta caaatgtacg gccagcaacg tcggttcgag 960
 atggcgctcg atgacgcca ctacctctga tagttgagtt gatacttcgg cgataaccgc 1020
 ttcacgagcc atggttcctt cctacaacg tcgagggggg gcccggtacc 1070

<210> 4

<211> 520

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: synthetic
 sequence coding for I-Ppo I homing-endonuclease

<220>

<221> CDS

<222> (16)..(504)

<400> 4

gcatatgaga tctcc atg gcg ctc acc aat gct caa atc ttg gct gtg att 51
 Met Ala Leu Thr Asn Ala Gln Ile Leu Ala Val Ile
 1 5 10

gac agt tgg gaa gaa aca gtc ggt cag ttt cca gtg ata acg cac cat 99
 Asp Ser Trp Glu Glu Thr Val Gly Gln Phe Pro Val Ile Thr His His
 15 20 25

gta cca tta ggt ggc ggt ctg caa gga acg ctc cat tgt tac gag atc 147
 Val Pro Leu Gly Gly Gly Leu Gln Gly Thr Leu His Cys Tyr Glu Ile
 30 35 40

ccc cta gca gct cct tat ggg gtt ggc ttt gct aag aat ggg cct acc 195
 Pro Leu Ala Ala Pro Tyr Gly Val Gly Phe Ala Lys Asn Gly Pro Thr
 45 50 55 60

cgc tgg caa tac aaa cgg aca atc aat caa gtc gtc cac aga tgg ggg 243
 Arg Trp Gln Tyr Lys Arg Thr Ile Asn Gln Val Val His Arg Trp Gly
 65 70 75

tcc cac aca gtc cct ttt cta tta gaa ccg gat aac atc aac ggc aaa 291
 Ser His Thr Val Pro Phe Leu Leu Glu Pro Asp Asn Ile Asn Gly Lys
 80 85 90

acc tgc aca gca tcg cac cta tgt cat aat act cga tgc cac aat ccc 339
 Thr Cys Thr Ala Ser His Leu Cys His Asn Thr Arg Cys His Asn Pro
 95 100 105

ttg cac ttg tgc tgg gag tca cta gac gac aac aaa ggc aga aac tgg 387
 Leu His Leu Cys Trp Glu Ser Leu Asp Asp Asn Lys Gly Arg Asn Trp
 110 115 120

tgc ccc ggt ccc aac ggg gga tgt gtc cat gcg gtg gtt tgt tta agg 435
 Cys Pro Gly Pro Asn Gly Gly Cys Val His Ala Val Val Cys Leu Arg
 125 130 135 140

cag ggt ccg ttg tac ggc cca ggg gcg act gtg gca ggt cct caa caa 483
 Gln Gly Pro Leu Tyr Gly Pro Gly Ala Thr Val Ala Gly Pro Gln Gln
 145 150 155

agg ggc agt cac ttt gtg gta taactgcagc tcgagg 520
 Arg Gly Ser His Phe Val Val
 160

<210> 5

<211> 163

<212> PRT

<213> Künstliche Sequenz

<223> Beschreibung der künstlichen Sequenz: synthetic
sequence coding for I-Ppo I homing-endonuclease

<400> 5

```

Met Ala Leu Thr Asn Ala Gln Ile Leu Ala Val Ile Asp Ser Trp Glu
  1             5             10             15
Glu Thr Val Gly Gln Phe Pro Val Ile Thr His His Val Pro Leu Gly
          20             25             30
Gly Gly Leu Gln Gly Thr Leu His Cys Tyr Glu Ile Pro Leu Ala Ala
          35             40             45
Pro Tyr Gly Val Gly Phe Ala Lys Asn Gly Pro Thr Arg Trp Gln Tyr
          50             55             60
Lys Arg Thr Ile Asn Gln Val Val His Arg Trp Gly Ser His Thr Val
          65             70             75             80
Pro Phe Leu Leu Glu Pro Asp Asn Ile Asn Gly Lys Thr Cys Thr Ala
          85             90             95
Ser His Leu Cys His Asn Thr Arg Cys His Asn Pro Leu His Leu Cys
          100            105            110
Trp Glu Ser Leu Asp Asp Asn Lys Gly Arg Asn Trp Cys Pro Gly Pro
          115            120            125
Asn Gly Gly Cys Val His Ala Val Val Cys Leu Arg Gln Gly Pro Leu
          130            135            140
Tyr Gly Pro Gly Ala Thr Val Ala Gly Pro Gln Gln Arg Gly Ser His
          145            150            155            160
Phe Val Val

```

<210> 6

<211> 21

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:
oligonucleotide primer

<400> 6

taaggccctc ggtagcaacg g

21

<210> 7

<211> 20

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:
oligonucleotide primer

<400> 7

ggggtaccaa atccaactag

20

<210> 8

<211> 25

<212> DNA

<213> Künstliche Sequenz

<220>
 <223> Beschreibung der künstlichen Sequenz:
 oligonucleotide primer
 <400> 8
 ggagctcgct cccccgccgt cggtc 25
 <210> 9
 <211> 27
 <212> DNA
 <213> Künstliche Sequenz
 <220>
 <223> Beschreibung der künstlichen Sequenz:
 oligonucleotide primer
 <400> 9
 gatgcatgat gacttgacgg catcctc 27
 <210> 10
 <211> 175
 <212> DNA
 <213> Pisum sativum
 <220>
 <221> CDS
 <222> (1)..(174)
 <223> plastid translocalization sequence
 <400> 10
 atg gct tct atg ata tcc tct tca gct gtg act aca gtc agc cgt gct 48
 Met Ala Ser Met Ile Ser Ser Ser Ala Val Thr Thr Val Ser Arg Ala
 1 5 10 15
 tct acg gtg caa tcg gcc gcg gtg gct cca ttc ggc ggc ctc aaa tcc 96
 Ser Thr Val Gln Ser Ala Ala Val Ala Pro Phe Gly Gly Leu Lys Ser
 20 25 30
 atg act gga ttc cca gtt aag aag gtc aac act gac att act tcc att 144
 Met Thr Gly Phe Pro Val Lys Lys Val Asn Thr Asp Ile Thr Ser Ile
 35 40 45
 aca agc aat ggt gga aga gta aag tgc atg c 175
 Thr Ser Asn Gly Gly Arg Val Lys Cys Met
 50 55
 <210> 11
 <211> 58
 <212> PRT
 <213> Pisum sativum
 <400> 11
 Met Ala Ser Met Ile Ser Ser Ser Ala Val Thr Thr Val Ser Arg Ala
 1 5 10 15
 Ser Thr Val Gln Ser Ala Ala Val Ala Pro Phe Gly Gly Leu Lys Ser
 20 25 30
 Met Thr Gly Phe Pro Val Lys Lys Val Asn Thr Asp Ile Thr Ser Ile
 35 40 45
 Thr Ser Asn Gly Gly Arg Val Lys Cys Met
 50 55
 <210> 12
 <211> 84

<212> PRT
 <213> *Nicotiana tabacum*
 <220>
 <221> TRANSIT
 <222> (1)..(84)
 <223> plastidic translocalisation signal derived from
 plastidic transketolase
 <400> 12
 Met Ala Ser Ser Ser Ser Leu Thr Leu Ser Gln Ala Ile Leu Ser Arg
 1 5 10 15
 Ser Val Pro Arg His Gly Ser Ala Ser Ser Ser Gln Leu Ser Pro Ser
 20 25 30
 Ser Leu Thr Phe Ser Gly Leu Lys Ser Asn Pro Asn Ile Thr Thr Ser
 35 40 45
 Arg Arg Arg Thr Pro Ser Ser Ala Ala Ala Ala Val Val Arg Ser
 50 55 60
 Pro Ala Ile Arg Ala Ser Ala Ala Thr Glu Thr Ile Glu Lys Thr Glu
 65 70 75 80
 Thr Ala Gly Ser

<210> 13
 <211> 258
 <212> DNA
 <213> *Nicotiana tabacum*
 <220>
 <221> transit_peptide
 <222> (1)..(258)
 <223> plastidic translocalisation signal derived from
 plastidic transketolase (frame 1)
 <400> 13
 ggtaccatgg cgtcttcttc ttctctcaact ctctctcaag ctatcctctc tcgttctgtc 60
 cctcgccatg gctctgcctc ttcttctcaa ctttcccctt cttctctcac tttttccggc 120
 cttaaatacca atcccaatat caccacctcc cgccgccgta ctccttcctc cgccgcccgc 180
 gccgcccgtc taaggtcacc ggcgattcgt gcctcagctg caaccgaaac catagagaaa 240
 actgagactg cgggatcc 258

<210> 14
 <211> 260
 <212> DNA
 <213> *Nicotiana tabacum*
 <220>
 <221> transit_peptide
 <222> (1)..(260)
 <223> plastidic translocalisation signal derived from
 plastidic transketolase (frame 2)

<400> 14
 ggtaccatgg cgtcttcttc ttctctcaact ctctctcaag ctatcctctc tcgttctgtc 60
 cctcgccatg gctctgcctc ttcttctcaa ctttcccctt cttctctcac tttttccggc 120
 cttaaatacca atcccaatat caccacctcc cgccgccgta ctccttcctc cgccgcccgc 180
 gccgcccgtc taaggtcacc ggcgattcgt gcctcagctg caaccgaaac catagagaaa 240
 actgagactg cgctggatcc 260

<210> 15
 <211> 259

```

<212> DNA
<213> Nicotiana tabacum
<220>
<221> transit_peptide
<222> (1)..(259)
<223> plastidic translocalisation signal derived from
        plastidic transketolase (frame 3)
<400> 15
ggtaccatgg cgtcttcttc ttctctcact ctctctcaag ctatcctctc tcgttctgtc 60
cctcgccatg gctctgcctc ttctctcaaa ctttcccctt cttctctcac tttttccggc 120
cttaaatacca atcccaatat caccacctcc cgccgccgta ctccttcctc cgccgccgcc 180
gccgccgtcg taaggtcacc ggcgattcgt gcctcagctg caaccgaaac catagagaaa 240
actgagactg cggggatcc                                     259
<210> 16
<211> 63
<212> PRT
<213> Arabidopsis thaliana
<220>
<221> TRANSIT
<222> (1)..(63)
<223> plastidic translocalisation signal derived from
        plastidic isopentenylpyrophosphate isomerase-2
        (IPP-2)
<400> 16
Met Ser Ala Ser Ser Leu Phe Asn Leu Pro Leu Ile Arg Leu Arg Ser
 1              5              10              15
Leu Ala Leu Ser Ser Phe Ser Ser Phe Arg Phe Ala His Arg Pro
          20              25              30
Leu Ser Ser Ile Ser Pro Arg Lys Leu Pro Asn Phe Arg Ala Phe Ser
          35              40              45
Gly Thr Ala Met Thr Asp Thr Lys Asp Gly Ser Arg Val Asp Met
          50              55              60
<210> 17
<211> 205
<212> DNA
<213> Arabidopsis thaliana
<220>
<221> transit_peptide
<222> (1)..(205)
<223> plastidic translocalisation signal derived from
        plastidic isopentenylpyrophosphate isomerase-2
        (IPP-2) (frame 1)
<400> 17
gatatccaca ccaacaccaa tgtctgcttc ttctttatct aatctcccat tgattcgctt 60
cagatctctc gctctttcgt cttctttttc ttctttccga tttgcccatc gtcctctgtc 120
atcgatttca ccgagaaagt taccgaatct tcgtgctttc tctggtaccg ctatgacaga 180
tactaaagat ggatcccggg tcgac                                     205
<210> 18
<211> 207
<212> DNA
<213> Arabidopsis thaliana
<220>
<221> transit_peptide

```

<222> (1)..(207)
<223> plastidic translocalisation signal derived from
plastidic isopentenylpyrophosphate isomerase-2
(IPP-2) (frame 2)

<400> 18
gatatccaca ccaacaccaa tgtctgcttc ttctttatctt aatctcccat tgattcgctt 60
cagatctctc gctctttcgt cttctttttc ttctttccga ttgcccac gtcctctgtc 120
atcgatttca ccgagaaagt taccgaattt tcgtgctttc tctggtaccg ctatgacaga 180
tactaaagat ctggatcccg ggtcgac 207

<210> 19
<211> 206
<212> DNA
<213> Arabidopsis thaliana

<220>
<221> transit_peptide
<222> (1)..(206)
<223> plastidic translocalisation signal derived from
plastidic isopentenylpyrophosphate isomerase-2
(IPP-2) (frame 3)

<400> 19
gatatccaca ccaacaccaa tgtctgcttc ttctttatctt aatctcccat tgattcgctt 60
cagatctctc gctctttcgt cttctttttc ttctttccga ttgcccac gtcctctgtc 120
atcgatttca ccgagaaagt taccgaattt tcgtgctttc tctggtaccg ctatgacaga 180
tactaaagat gggatcccg gtcgac 206

<210> 20
<211> 44
<212> DNA
<213> Nicotiana tabacum

<220>
<221> promoter
<222> (1)..(44)
<223> plastidic promoter PrbcL

<400> 20
gttgcgctat atatatgaaa gagtatacaa taatgatgta ttgt 44

<210> 21
<211> 64
<212> DNA
<213> Nicotiana tabacum

<220>
<221> promoter
<222> (1)..(64)
<223> plastidic promoter Prps16-107

<400> 21
tagcgatggg gtcttactaa agaaaaatat ttatccacct atctctatag tatatagata 60
taga 64

<210> 22
<211> 91
<212> DNA
<213> Nicotiana tabacum

<220>
<221> promoter
<222> (1)..(91)
<223> plastidic promoter Prn16

<400> 22
 cgccgtcgtt caatgagaat ggataagagg ctccgtgggat tgacgtgagg gggcagggat 60
 ggctatatatt ctgggagcga actccgggcg a 91

<210> 23
 <211> 71
 <212> DNA
 <213> Arabidopsis thaliana

<220>
 <221> promoter
 <222> (1)..(71)
 <223> plastidic promoter PaccD-129

<400> 23
 gtcgacatat tatttttaaat aatataaagg ggggtccaac atattaatat atagtgaagt 60
 gttccggatc c 71

<210> 24
 <211> 27
 <212> DNA
 <213> Nicotiana tabacum

<220>
 <221> promoter
 <222> (1)..(27)
 <223> plastidic promoter PclpP-53

<400> 24
 agacaataaa aaaaattggt acgtttc 27

<210> 25
 <211> 65
 <212> DNA
 <213> Nicotiana tabacum

<220>
 <221> promoter
 <222> (1)..(65)
 <223> plastidic promoter Prn-62

<400> 25
 gagcgaactc cgggcgaata tgaagcgcat ggatacaagt tatgccttgg aatgaaagac 60
 aattc 65

<210> 26
 <211> 107
 <212> DNA
 <213> Nicotiana tabacum

<220>
 <221> promoter
 <222> (1)..(107)
 <223> plastidic promoter Prps16

<400> 26
 tctatatata tttatatatt aaatattaat taatatgtta gataaattaa tattagcgat 60
 ggggtcttac taaagaaaaa tatttatcca cctatctcta tagtata 107

<210> 27
 <211> 47
 <212> DNA
 <213> Nicotiana tabacum

<220>
 <221> promoter

<222> (1)..(47)
 <223> plastidic promoter PatpB/E-290
 <400> 27
 agaaatagaa aataaagttc aggttcgaat tccatagaat agataat 47
 <210> 28
 <211> 25
 <212> DNA
 <213> Nicotiana tabacum
 <220>
 <221> promoter
 <222> (1)..(25)
 <223> plastidic promoter PrpoB-345
 <400> 28
 aatgtgtatt atcataataa tggta 25
 <210> 29
 <211> 35
 <212> DNA
 <213> Künstliche Sequenz
 <220>
 <223> Beschreibung der künstlichen Sequenz: promoter
 sequence derived from consensus sequence of
 E.coli sigma70 promoters
 <400> 29
 ttgacattca ctcttcaatt atctataatg ataca 35
 <210> 30
 <211> 80
 <212> DNA
 <213> Nicotiana tabacum
 <220>
 <221> misc_feature
 <222> (1)..(80)
 <223> 5'-untranslated region of psbA gene
 <400> 30
 tccattttct attttgattt gtagaaaact agtgtgcttg ggagtccttg atgattaaat 60
 aaaccaagat ttaccatgg 80
 <210> 31
 <211> 40
 <212> DNA
 <213> Nicotiana tabacum
 <220>
 <221> misc_feature
 <222> (1)..(40)
 <223> 5'-untranslated region and part of translated
 region of rbcL gene
 <400> 31
 agttgtaggg agggattcat gacaccacaa acagccatgg 40
 <210> 32
 <211> 22
 <212> DNA
 <213> Nicotiana tabacum
 <220>
 <221> misc_feature

<222> (1)..(22)
 <223> 5'-untranslated region of rbcL gene
 <400> 32
 agttgtaggg agggattcat ga 22
 <210> 33
 <211> 127
 <212> DNA
 <213> Synechocystis sp.
 <220>
 <221> misc_feature
 <222> (1)..(127)
 <223> 3'-untranslated region of Synechocystis psbA-1 gene
 <400> 33
 tgccattgcc ataactgctt tcggttagac ttcgtttcat ttggttaatc aagggcactc 60
 tcgcaatggg gtgcctttta tggccaagg ttaaagttaa gccagtacta tttctagggt 120
 gaaatgt 127
 <210> 34
 <211> 90
 <212> DNA
 <213> Nicotiana tabacum
 <220>
 <221> misc_feature
 <222> (1)..(90)
 <223> 3'-untranslated region of tobacco psbA gene
 <400> 34
 cctggcctag tctataggag gttttgaaaa gaaaggagca ataatcattt tcttgttcta 60
 tcaagagggt gctattgctc ctttcttttt 90
 <210> 35
 <211> 136
 <212> DNA
 <213> Nicotiana tabacum
 <220>
 <221> misc_feature
 <222> (1)..(136)
 <223> 3'-untranslated region of tobacco rbcL gene
 <400> 35
 agtagacatt agcagataaa ttagcaggaa ataaagaagg ataaggagaa agaactcaag 60
 taattatcct tcgttctctt aattgaattg caattaaact cggccaatc ttttactaaa 120
 aggattgagc cgaata 136
 <210> 36
 <211> 13
 <212> DNA
 <213> Künstliche Sequenz
 <220>
 <223> Beschreibung der künstlichen Sequenz: synthetic
 ribosome binding site (RBS)
 <400> 36
 ggaggnnnnn atg 13
 <210> 37
 <211> 13
 <212> DNA
 <213> Künstliche Sequenz

<220>
 <223> Beschreibung der künstlichen Sequenz: synthetic
 ribosome binding site (RBS)
 <400> 37
 ggaggatctc atg 13
 <210> 38
 <211> 34
 <212> DNA
 <213> Künstliche Sequenz
 <220>
 <223> Beschreibung der künstlichen Sequenz:
 oligonucleotide primer
 <400> 38
 atggatccat atggccatgg cacaaggggt tgtg 34
 <210> 39
 <211> 33
 <212> DNA
 <213> Künstliche Sequenz
 <220>
 <223> Beschreibung der künstlichen Sequenz:
 oligonucleotide primer
 <400> 39
 cccgggctcg agctgcagct acgccgctac gtc 33
 <210> 40
 <211> 2815
 <212> DNA
 <213> Künstliche Sequenz
 <220>
 <223> Beschreibung der künstlichen Sequenz: insert of
 binary vector pCB193-17
 <220>
 <221> CDS
 <222> (749)..(2587)
 <223> ORF for Phi-C31 recombinase
 <220>
 <221> CDS
 <222> (563)..(2587)
 <223> ORF for fusion protein of Phi-C31 recombinase and
 plastidic transit peptide
 <400> 40
 gaattccatg gagtcaaaga ttcaaataga ggacctaaca gaactcgccg taaagactgg 60
 cgaacagttc atacagagtc tcttacgact caatgacaag aagaaaatct tcgtcaacat 120
 ggtggagcac gacacgcttg tctactccaa aaatatcaaa gatacagtct cagaagacca 180
 aagggcaatt gagacttttc aacaaaggggt aatatccgga aacctcctcg gattccattg 240
 cccagctatc tgtcacttta ttgtgaagat agtggaaaag gaaggtggct cctacaaatg 300
 ccatcattgc gataaaagaa aggccatcgt tgaagatgcc tctgccgaca gtggtcccaa 360
 agatggaccc ccaccacga ggagcatcgt ggaaaaagaa gacgttccaa ccacgtcttc 420
 aaagcaagtg gattgatgtg atatctccac tgacgtaagg gatgacgcac aatcccacta 480
 tccttcgcaa gacccttcct ctatataagg aagttcattt catttgaggaga ggacagggta 540

cccgatatcc	acaaccaaca	cc	atg	tct	gct	tct	tct	tct	tta	ttt	aat	ctc	cca	592
			Met	Ser	Ala	Ser	Ser	Ser	Leu	Phe	Asn	Leu	Pro	
			1					5					10	
ttg	att	cgc	ctc	aga	tct	ctc	gct	ctt	tcg	tct	tct	ttt	tct	640
Leu	Ile	Arg	Leu	Arg	Ser	Leu	Ala	Leu	Ser	Ser	Ser	Phe	Ser	
				15					20				25	
cga	ttt	gcc	cat	cgt	cct	ctg	tca	tcg	att	tca	ccg	aga	aag	688
Arg	Phe	Ala	His	Arg	Pro	Leu	Ser	Ser	Ile	Ser	Pro	Arg	Lys	
			30					35					40	
aat	ttt	cgt	gct	ttc	tct	ggc	acc	gct	atg	aca	gat	act	aaa	736
Asn	Phe	Arg	Ala	Phe	Ser	Gly	Thr	Ala	Met	Thr	Asp	Thr	Lys	
		45					50					55		
atc	cat	atg	gcc	atg	gca	caa	ggg	gtt	gtg	acc	ggg	gtg	gac	784
Ile	His	Met	Ala	Met	Ala	Gln	Gly	Val	Val	Thr	Gly	Val	Asp	
	60					65					70			
gcg	ggc	gct	tac	gac	cgt	cag	tcg	cgc	gag	cgc	gag	aat	tcg	832
Ala	Gly	Ala	Tyr	Asp	Arg	Gln	Ser	Arg	Glu	Arg	Glu	Asn	Ser	
75					80					85				
gca	agc	cca	gcg	aca	cag	cgt	agc	gcc	aac	gaa	gac	aag	gcg	880
Ala	Ser	Pro	Ala	Thr	Gln	Arg	Ser	Ala	Asn	Glu	Asp	Lys	Ala	
			95						100				105	
ctt	cag	cgc	gaa	gtc	gag	cgc	gac	ggg	ggc	cgg	ttc	agg	ttc	928
Leu	Gln	Arg	Glu	Val	Glu	Arg	Asp	Gly	Gly	Arg	Phe	Arg	Phe	
			110					115					120	
cat	ttc	agc	gaa	gcg	ccg	ggc	acg	tcg	gcg	ttc	ggg	acg	gcg	976
His	Phe	Ser	Glu	Ala	Pro	Gly	Thr	Ser	Ala	Phe	Gly	Thr	Ala	
		125					130					135		
ccg	gag	ttc	gaa	cgc	atc	ctg	aac	gaa	tgc	cgc	gcc	ggg	cgg	1024
Pro	Glu	Phe	Glu	Arg	Ile	Leu	Asn	Glu	Cys	Arg	Ala	Gly	Arg	
	140					145					150			
atg	atc	att	gtc	tat	gac	gtg	tcg	cgc	ttc	tcg	cgc	ctg	aag	1072
Met	Ile	Ile	Val	Tyr	Asp	Val	Ser	Arg	Phe	Ser	Arg	Leu	Lys	
155					160					165				
gac	gcg	att	ccg	att	gtc	tcg	gaa	ttg	ctc	gcc	ctg	ggc	gtg	1120
Asp	Ala	Ile	Pro	Ile	Val	Ser	Glu	Leu	Leu	Ala	Leu	Gly	Val	
			175						180				185	
gtt	tcc	act	cag	gaa	ggc	gtc	ttc	cgg	cag	gga	aac	gtc	atg	1168
Val	Ser	Thr	Gln	Glu	Gly	Val	Phe	Arg	Gln	Gly	Asn	Val	Met	
			190					195					200	
att	cac	ctg	att	atg	cgg	ctc	gac	gcg	tcg	cac	aaa	gaa	tct	1216
Ile	His	Leu	Ile	Met	Arg	Leu	Asp	Ala	Ser	His	Lys	Glu	Ser	
		205					210					215		
aag	tcg	gcg	aag	att	ctc	gac	acg	aag	aac	ctt	cag	cgc	gaa	1264
Lys	Ser	Ala	Lys	Ile	Leu	Asp	Thr	Lys	Asn	Leu	Gln	Arg	Glu	
		220				225								

15

aag	ctt	gcg	cac	tcg	acc	act	ccc	ctt	acc	gga	ccc	ttc	gag	ttc	gag	1408
Lys	Leu	Ala	His	Ser	Thr	Thr	Pro	Leu	Thr	Gly	Pro	Phe	Glu	Phe	Glu	
		270						275					280			
ccc	gac	gta	atc	cgg	tgg	tgg	tgg	cgt	gag	atc	aag	acg	cac	aaa	cac	1456
Pro	Asp	Val	Ile	Arg	Trp	Trp	Trp	Arg	Glu	Ile	Lys	Thr	His	Lys	His	
		285						290					295			
ctt	ccc	ttc	aag	cgg	ggc	agt	caa	gcc	gcc	att	cac	ccg	ggc	agc	atc	1504
Leu	Pro	Phe	Lys	Pro	Gly	Ser	Gln	Ala	Ala	Ile	His	Pro	Gly	Ser	Ile	
	300					305					310					
acg	ggg	ctt	tgt	aag	cgc	atg	gac	gct	gac	gcc	gtg	ccg	acc	cgg	ggc	1552
Thr	Gly	Leu	Cys	Lys	Arg	Met	Asp	Ala	Asp	Ala	Val	Pro	Thr	Arg	Gly	
315					320					325					330	
gag	acg	att	ggg	aag	aag	acc	gct	tca	agc	gcc	tgg	gac	ccg	gca	acc	1600
Glu	Thr	Ile	Gly	Lys	Lys	Thr	Ala	Ser	Ser	Ala	Trp	Asp	Pro	Ala	Thr	
			335						340					345		
gtt	atg	cga	atc	ctt	cgg	gac	ccg	cgt	att	gcg	ggc	ttc	gcc	gct	gag	1648
Val	Met	Arg	Ile	Leu	Arg	Asp	Pro	Arg	Ile	Ala	Gly	Phe	Ala	Ala	Glu	
		350						355					360			
gtg	atc	tac	aag	aag	aag	ccg	gac	ggc	acg	ccg	acc	acg	aag	att	gag	1696
Val	Ile	Tyr	Lys	Lys	Lys	Pro	Asp	Gly	Thr	Pro	Thr	Thr	Lys	Ile	Glu	
		365					370					375				
ggt	tac	cgc	att	cag	cgc	gac	ccg	atc	acg	ctc	cgg	ccg	gtc	gag	ctt	1744
Gly	Tyr	Arg	Ile	Gln	Arg	Asp	Pro	Ile	Thr	Leu	Arg	Pro	Val	Glu	Leu	
	380					385					390					
gat	tgc	gga	ccg	atc	atc	gag	ccc	gct	gag	tgg	tat	gag	ctt	cag	gcg	1792
Asp	Cys	Gly	Pro	Ile	Ile	Glu	Pro	Ala	Glu	Trp	Tyr	Glu	Leu	Gln	Ala	
395					400					405					410	
tgg	ttg	gac	ggc	agg	ggg	cgc	ggc	aag	ggg	ctt	tcc	cgg	ggg	caa	gcc	1840
Trp	Leu	Asp	Gly	Arg	Gly	Arg	Gly	Lys	Gly	Leu	Ser	Arg	Gly	Gln	Ala	
			415						420					425		
att	ctg	tcc	gcc	atg	gac	aag	ctg	tac	tgc	gag	tgt	ggc	gcc	gtc	atg	1888
Ile	Leu	Ser	Ala	Met	Asp	Lys	Leu	Tyr	Cys	Glu	Cys	Gly	Ala	Val	Met	
		430						435					440			
act	tgc	aag	cgc	ggg	gaa	gaa	tgc	atc	aag	gac	tct	tac	cgc	tgc	cgt	1936
Thr	Ser	Lys	Arg	Gly	Glu	Glu	Ser	Ile	Lys	Asp	Ser	Tyr	Arg	Cys	Arg	
		445					450					455				
cgc	cgg	aag	gtg	gtc	gac	ccg	tcc	gca	cct	ggg	cag	cac	gaa	ggc	acg	1984
Arg	Arg	Lys	Val	Val	Asp	Pro	Ser	Ala	Pro	Gly	Gln	His	Glu	Gly	Thr	
		460				465					470					
tgc	aac	gtc	agc	atg	gcg	gca	ctc	gac	aag	ttc	gtt	gcg	gaa	cgc	atc	2032
Cys	Asn	Val	Ser	Met	Ala	Ala	Leu	Asp	Lys	Phe	Val	Ala	Glu	Arg	Ile	
475					480					485					490	
ttc	aac	aag	atc	agg	cac	gcc	gaa	ggc	gac	gaa	gag	acg	ttg	gcg	ctt	2080
Phe	Asn	Lys	Ile	Arg	His	Ala	Glu	Gly	Asp	Glu	Glu	Thr	Leu	Ala	Leu	
			495						500					505		
ctg	tgg	gaa	gcc	gcc	cga	cgc	ttc	ggc	aag	ctc	act	gag	gcg	cct	gag	2128
Leu	Trp	Glu	Ala	Ala	Arg	Arg	Phe	Gly	Lys	Leu	Thr	Glu	Ala	Pro	Glu	
		510						515					520			
aag	agc	ggc	gaa	cgg	gcg	aac	ctt	gtt	gcg	gag	cgc	gcc	gac	gcc	ctg	2176
Lys	Ser	Gly	Glu	Arg	Ala	Asn	Leu	Val	Ala	Glu	Arg	Ala	Asp	Ala	Leu	
		525					530					535				

16

aac gcc ctt gaa gag ctg tac gaa gac cgc gcg gca ggc gcg tac gac 2224
 Asn Ala Leu Glu Glu Leu Tyr Glu Asp Arg Ala Ala Gly Ala Tyr Asp
 540 545 550

gga ccc gtt ggc agg aag cac ttc cgg aag caa cag gca gcg ctg acg 2272
 Gly Pro Val Gly Arg Lys His Phe Arg Lys Gln Gln Ala Ala Leu Thr
 555 560 565 570

ctc cgg cag caa ggg gcg gaa gag cgg ctt gcc gaa ctt gaa gcc gcc 2320
 Leu Arg Gln Gln Gly Ala Glu Glu Arg Leu Ala Glu Leu Glu Ala Ala
 575 580 585

gaa gcc ccg aag ctt ccc ctt gac caa tgg ttc ccc gaa gac gcc gac 2368
 Glu Ala Pro Lys Leu Pro Leu Asp Gln Trp Phe Pro Glu Asp Ala Asp
 590 595 600

gct gac ccg acc ggc cct aag tcg tgg tgg ggg cgc gcg tca gta gac 2416
 Ala Asp Pro Thr Gly Pro Lys Ser Trp Trp Gly Arg Ala Ser Val Asp
 605 610 615

gac aag cgc gtg ttc gtc ggg ctc ttc gta gac aag atc gtt gtc acg 2464
 Asp Lys Arg Val Phe Val Gly Leu Phe Val Asp Lys Ile Val Val Thr
 620 625 630

aag tcg act acg ggc agg ggg cag gga acg ccc atc gag aag cgc gct 2512
 Lys Ser Thr Thr Gly Arg Gly Gln Gly Thr Pro Ile Glu Lys Arg Ala
 635 640 645 650

tcg atc acg tgg gcg aag ccg ccg acc gac gac gac gaa gac gac gcc 2560
 Ser Ile Thr Trp Ala Lys Pro Pro Thr Asp Asp Asp Glu Asp Asp Ala
 655 660 665

cag gac ggc acg gaa gac gta gcg gcg tagctgcagc tcgacctgca 2607
 Gln Asp Gly Thr Glu Asp Val Ala Ala
 670 675

ggcatgccct gctttaatga gatatgcgag acgcctatga tcgcatgata tttgctttca 2667
 attctgttgc gcacgttgta aaaaacctga gcatgtgtag ctcagatcct tacccgccgt 2727
 ttccggttcat tctaataaat atatcaccgc ttactatcgt atttttatga ataataattct 2787
 ccgttcaatt tactgattgt ccaagctt 2815

<210> 41
 <211> 675
 <212> PRT
 <213> Künstliche Sequenz
 <223> Beschreibung der künstlichen Sequenz: insert of
 binary vector pCB193-17

<400> 41
 Met Ser Ala Ser Ser Leu Phe Asn Leu Pro Leu Ile Arg Leu Arg Ser
 1 5 10 15
 Leu Ala Leu Ser Ser Phe Ser Ser Phe Arg Phe Ala His Arg Pro
 20 25 30
 Leu Ser Ser Ile Ser Pro Arg Lys Leu Pro Asn Phe Arg Ala Phe Ser
 35 40 45
 Gly Thr Ala Met Thr Asp Thr Lys Asp Gly Ile His Met Ala Met Ala
 50 55 60
 Gln Gly Val Val Thr Gly Val Asp Thr Tyr Ala Gly Ala Tyr Asp Arg
 65 70 75 80
 Gln Ser Arg Glu Arg Glu Asn Ser Ser Ala Ala Ser Pro Ala Thr Gln
 85 90 95
 Arg Ser Ala Asn Glu Asp Lys Ala Ala Asp Leu Gln Arg Glu Val Glu

			100					105					110				
Arg	Asp	Gly	Gly	Arg	Phe	Arg	Phe	Val	Gly	His	Phe	Ser	Glu	Ala	Pro		
		115					120					125					
Gly	Thr	Ser	Ala	Phe	Gly	Thr	Ala	Glu	Arg	Pro	Glu	Phe	Glu	Arg	Ile		
		130				135					140						
Leu	Asn	Glu	Cys	Arg	Ala	Gly	Arg	Leu	Asn	Met	Ile	Ile	Val	Tyr	Asp		
145				150						155					160		
Val	Ser	Arg	Phe	Ser	Arg	Leu	Lys	Val	Met	Asp	Ala	Ile	Pro	Ile	Val		
			165						170					175			
Ser	Glu	Leu	Leu	Ala	Leu	Gly	Val	Thr	Ile	Val	Ser	Thr	Gln	Glu	Gly		
			180					185					190				
Val	Phe	Arg	Gln	Gly	Asn	Val	Met	Asp	Leu	Ile	His	Leu	Ile	Met	Arg		
		195				200						205					
Leu	Asp	Ala	Ser	His	Lys	Glu	Ser	Ser	Leu	Lys	Ser	Ala	Lys	Ile	Leu		
		210			215						220						
Asp	Thr	Lys	Asn	Leu	Gln	Arg	Glu	Leu	Gly	Gly	Tyr	Val	Gly	Gly	Lys		
225				230					235						240		
Ala	Pro	Tyr	Gly	Phe	Glu	Leu	Val	Ser	Glu	Thr	Lys	Glu	Ile	Thr	Arg		
			245					250						255			
Asn	Gly	Arg	Met	Val	Asn	Val	Val	Ile	Asn	Lys	Leu	Ala	His	Ser	Thr		
			260					265					270				
Thr	Pro	Leu	Thr	Gly	Pro	Phe	Glu	Phe	Glu	Pro	Asp	Val	Ile	Arg	Trp		
		275					280					285					
Trp	Trp	Arg	Glu	Ile	Lys	Thr	His	Lys	His	Leu	Pro	Phe	Lys	Pro	Gly		
		290			295						300						
Ser	Gln	Ala	Ala	Ile	His	Pro	Gly	Ser	Ile	Thr	Gly	Leu	Cys	Lys	Arg		
305				310						315					320		
Met	Asp	Ala	Asp	Ala	Val	Pro	Thr	Arg	Gly	Glu	Thr	Ile	Gly	Lys	Lys		
			325					330						335			
Thr	Ala	Ser	Ser	Ala	Trp	Asp	Pro	Ala	Thr	Val	Met	Arg	Ile	Leu	Arg		
			340					345					350				
Asp	Pro	Arg	Ile	Ala	Gly	Phe	Ala	Ala	Glu	Val	Ile	Tyr	Lys	Lys	Lys		
		355					360					365					
Pro	Asp	Gly	Thr	Pro	Thr	Thr	Lys	Ile	Glu	Gly	Tyr	Arg	Ile	Gln	Arg		
		370			375						380						
Asp	Pro	Ile	Thr	Leu	Arg	Pro	Val	Glu	Leu	Asp	Cys	Gly	Pro	Ile	Ile		
385				390						395					400		
Glu	Pro	Ala	Glu	Trp	Tyr	Glu	Leu	Gln	Ala	Trp	Leu	Asp	Gly	Arg	Gly		
			405					410						415			
Arg	Gly	Lys	Gly	Leu													

[illegible]

<223> Beschreibung der künstlichen Sequenz: insert of
binary vector pCB193-17

Met	Ala	Gln	Gly	Val	Val	Thr	Gly	Val	Asp	Thr	Tyr	Ala	Gly	Ala	Tyr
1				5					10					15	
Asp	Arg	Gln	Ser	Arg	Glu	Arg	Glu	Asn	Ser	Ser	Ala	Ala	Ser	Pro	Ala
			20					25					30		
Thr	Gln	Arg	Ser	Ala	Asn	Glu	Asp	Lys	Ala	Ala	Asp	Leu	Gln	Arg	Glu
		35					40					45			
Val	Glu	Arg	Asp	Gly	Gly	Arg	Phe	Arg	Phe	Val	Gly	His	Phe	Ser	Glu
	50					55					60				
Ala	Pro	Gly	Thr	Ser	Ala	Phe	Gly	Thr	Ala	Glu	Arg	Pro	Glu	Phe	Glu
65					70					75					80
Arg	Ile	Leu	Asn	Glu	Cys	Arg	Ala	Gly	Arg	Leu	Asn	Met	Ile	Ile	Val
			85						90				95		
Tyr	Asp	Val	Ser	Arg	Phe	Ser	Arg	Leu	Lys	Val	Met	Asp	Ala	Ile	Pro
			100					105					110		
Ile	Val	Ser	Glu	Leu	Leu	Ala	Leu	Gly	Val	Thr	Ile	Val	Ser	Thr	Gln
		115					120					125			
Glu	Gly	Val	Phe	Arg	Gln	Gly	Asn	Val	Met	Asp	Leu	Ile	His	Leu	Ile
	130					135					140				
Met	Arg	Leu	Asp	Ala	Ser	His	Lys	Glu	Ser	Ser	Leu	Lys	Ser	Ala	Lys
145					150					155				160	
Ile	Leu	Asp	Thr	Lys	Asn	Leu	Gln	Arg	Glu	Leu	Gly	Gly	Tyr	Val	Gly
			165						170				175		
Gly	Lys	Ala	Pro	Tyr	Gly	Phe	Glu	Leu	Val	Ser	Glu	Thr	Lys	Glu	Ile
			180					185					190		
Thr	Arg	Asn	Gly	Arg	Met	Val	Asn	Val	Val	Ile	Asn	Lys	Leu	Ala	His
		195				200						205			
Ser	Thr	Thr	Pro	Leu	Thr	Gly	Pro	Phe	Glu	Phe	Glu	Pro	Asp	Val	Ile
	210					215					220				
Arg	Trp	Trp	Trp	Arg	Glu	Ile	Lys	Thr	His	Lys	His	Leu	Pro	Phe	Lys
225					230					235				240	
Pro	Gly	Ser	Gln	Ala	Ala	Ile	His	Pro	Gly	Ser	Ile	Thr	Gly	Leu	Cys
			245						250				255		
Lys	Arg	Met	Asp	Ala	Asp	Ala	Val	Pro	Thr	Arg	Gly	Glu	Thr	Ile	Gly

19

260							265					270				
Lys	Lys	Thr	Ala	Ser	Ser	Ala	Trp	Asp	Pro	Ala	Thr	Val	Met	Arg	Ile	
		275					280					285				
Leu	Arg	Asp	Pro	Arg	Ile	Ala	Gly	Phe	Ala	Ala	Glu	Val	Ile	Tyr	Lys	
		290				295					300					
Lys	Lys	Pro	Asp	Gly	Thr	Pro	Thr	Thr	Lys	Ile	Glu	Gly	Tyr	Arg	Ile	
305					310					315					320	
Gln	Arg	Asp	Pro	Ile	Thr	Leu	Arg	Pro	Val	Glu	Leu	Asp	Cys	Gly	Pro	
				325					330					335		
Ile	Ile	Glu	Pro	Ala	Glu	Trp	Tyr	Glu	Leu	Gln	Ala	Trp	Leu	Asp	Gly	
			340					345					350			
Arg	Gly	Arg	Gly	Lys	Gly	Leu	Ser	Arg	Gly	Gln	Ala	Ile	Leu	Ser	Ala	
			355					360				365				
Met	Asp	Lys	Leu	Tyr	Cys	Glu	Cys	Gly	Ala	Val	Met	Thr	Ser	Lys	Arg	
		370				375					380					
Gly	Glu	Glu	Ser	Ile	Lys	Asp	Ser	Tyr	Arg	Cys	Arg	Arg	Arg	Lys	Val	
385					390					395					400	
Val	Asp	Pro	Ser	Ala	Pro	Gly	Gln	His	Glu	Gly	Thr	Cys	Asn	Val	Ser	
				405					410					415		
Met	Ala	Ala	Leu	Asp	Lys	Phe	Val	Ala	Glu	Arg	Ile	Phe	Asn	Lys	Ile	
			420					425					430			
Arg	His	Ala	Glu	Gly	Asp	Glu	Glu	Thr	Leu	Ala	Leu	Leu	Trp	Glu	Ala	
		435						440					445			
Ala	Arg	Arg	Phe	Gly	Lys	Leu	Thr	Glu	Ala	Pro	Glu	Lys	Ser	Gly	Glu	
		450				455					460					
Arg	Ala	Asn	Leu	Val	Ala	Glu	Arg	Ala	Asp	Ala	Leu	Asn	Ala	Leu	Glu	
465					470					475					480	
Glu	Leu	Tyr	Glu	Asp	Arg	Ala	Ala	Gly	Ala	Tyr	Asp	Gly	Pro	Val	Gly	
				485				490					495			
Arg	Lys	His	Phe	Arg	Lys	Gln	Gln	Ala	Ala	Leu	Thr	Leu	Arg	Gln	Gln	
			500					505					510			
Gly	Ala	Glu	Glu	Arg	Leu	Ala	Glu	Leu	Glu	Ala	Ala	Glu	Ala	Pro	Lys	
		515						520				525				
Leu	Pro	Leu	Asp	Gln	Trp	Phe	Pro	Glu	Asp	Ala	Asp	Ala	Asp	Pro	Thr	
		530				535					540					
Gly	Pro	Lys	Ser	Trp	Trp	Gly	Arg	Ala	Ser	Val	Asp	Asp	Lys	Arg	Val	
545					550					555					560	
Phe	Val	Gly	Leu	Phe	Val	Asp	Lys	Ile	Val	Val	Thr	Lys	Ser	Thr	Thr	
				565					570					575		
Gly	Arg	Gly	Gln	Gly	Thr	Pro	Ile	Glu	Lys	Arg	Ala	Ser	Ile	Thr	Trp	
			580					585					590			
Ala	Lys	Pro	Pro	Thr	Asp	Asp	Asp	Glu	Asp	Asp	Ala	Gln	Asp	Gly	Thr	

<210> 43

<211> 31

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:
oligonucleotide primer

<400> 43

tggatccata tggccatggc taagaaagta g

<210> 44
 <211> 26
 <212> DNA
 <213> Künstliche Sequenz

<220>
 <223> Beschreibung der künstlichen Sequenz:
 oligonucleotide primer

<400> 44
 cctcgagctg cagttaagcg agttgg

26

<210> 45
 <211> 2430
 <212> DNA
 <213> Künstliche Sequenz

<220>
 <223> Beschreibung der künstlichen Sequenz: insert of
 binary vector pCB245-37

<220>
 <221> CDS
 <222> (748)..(2202)
 <223> ORF for TP901-1 recombinase

<220>
 <221> CDS
 <222> (568)..(2202)
 <223> ORF for fusionprotein of TP901-1 recombinase and
 plastidic transit peptide

<400> 45
 tgaattccat gcatggagtc aaagattcaa atagaggacc taacagaact cgccgtaaag 60
 actggcgaaac agttcatata gactctctta cgactcaatg acaagaagaa aatcttcgtc 120
 aacatgggtgg agcacgacac gcttgtctac tccaaaaata tcaaagatac agtctcagaa 180
 gaccaaaggg caattgagac ttttcaacaa agggtaatat ccggaaacct cctcggattc 240
 cattgcccag ctatctgtca ctttattgtg aagatagtgg aaaaggaagg tggctcctac 300
 aaatgccatc attgcgataa aagaaaggcc atcgttgaag atgcctctgc cgacagtggg 360
 cccaaagatg gacccccacc cacgaggagc atcgtaggaaa aagaagacgt tccaaccacg 420
 tcttcaaagc aagtggattg atgtgatata tccactgacg taagggatga cgcacaatcc 480
 cactatcctt cgcaagaccc ttcctctata taaggaagtt catttcattt ggagaggaca 540
 gggtagccga tatccacaac caacacc atg tct gct tct tct tta ttt aat ctc 594
 Met Ser Ala Ser Ser Leu Phe Asn Leu

1

5

cca ttg att cgc ctc aga tct ctc gct ctt tcg tct tct ttt tct tct 642
 Pro Leu Ile Arg Leu Arg Ser Leu Ala Leu Ser Ser Ser Phe Ser Ser
 10 15 20 25

ttc cga ttt gcc cat cgt cct ctg tca tcg att tca ccg aga aag tta 690
 Phe Arg Phe Ala His Arg Pro Leu Ser Ser Ile Ser Pro Arg Lys Leu
 30 35 40

ccg aat ttt cgt gct ttc tct ggt acc gct atg aca gat act aaa gat 738
 Pro Asn Phe Arg Ala Phe Ser Gly Thr Ala Met Thr Asp Thr Lys Asp
 45 50 55

ggg atc ccc atg gct aag aaa gca gca atc tat aca cga gta tcc act 786
 Gly Ile Pro Met Ala Lys Lys Ala Ala Ile Tyr Thr Arg Val Ser Thr
 60 65 70

21

act aac caa gca gag gaa ggg ttc tca att gat gag caa att gac cgt	834
Thr Asn Gln Ala Glu Glu Gly Phe Ser Ile Asp Glu Gln Ile Asp Arg	
75 80 85	
tta aca aaa tat gct gaa gca atg ggg tgg caa gta tct gat act tat	882
Leu Thr Lys Tyr Ala Glu Ala Met Gly Trp Gln Val Ser Asp Thr Tyr	
90 95 100 105	
act gat gct ggt ttt tca ggg gcc aaa ctt gaa cgc cca gca atg caa	930
Thr Asp Ala Gly Phe Ser Gly Ala Lys Leu Glu Arg Pro Ala Met Gln	
110 115 120	
aga tta atc aac gat atc gag aat aaa gct ttt gat aca gtt ctt gta	978
Arg Leu Ile Asn Asp Ile Glu Asn Lys Ala Phe Asp Thr Val Leu Val	
125 130 135	
tat aag cta gac cgc ctt tca cgt agt gta aga gat act ctt tat ctt	1026
Tyr Lys Leu Asp Arg Leu Ser Arg Ser Val Arg Asp Thr Leu Tyr Leu	
140 145 150	
gtt aag gat gtg ttc aca aaa aat aaa ata gac ttt atc tcg ctt aat	1074
Val Lys Asp Val Phe Thr Lys Asn Lys Ile Asp Phe Ile Ser Leu Asn	
155 160 165	
gaa agt att gat act tct tct gct atg ggt agc ttg ttt ctc act att	1122
Glu Ser Ile Asp Thr Ser Ser Ala Met Gly Ser Leu Phe Leu Thr Ile	
170 175 180 185	
ctt tct gca att aat gag ttt gaa aga gag aat ata aaa gaa cgc atg	1170
Leu Ser Ala Ile Asn Glu Phe Glu Arg Glu Asn Ile Lys Glu Arg Met	
190 195 200	
act atg ggt aaa cta ggg cga gcg aaa tct ggt aag tct atg atg tgg	1218
Thr Met Gly Lys Leu Gly Arg Ala Lys Ser Gly Lys Ser Met Met Trp	
205 210 215	
act aag aca gct ttt ggg tat tac cac aac aga aag aca ggt ata tta	1266
Thr Lys Thr Ala Phe Gly Tyr Tyr His Asn Arg Lys Thr Gly Ile Leu	
220 225 230	
gaa att gtt cct tta caa gct aca ata gtt gaa caa ata ttc act gat	1314
Glu Ile Val Pro Leu Gln Ala Thr Ile Val Glu Gln Ile Phe Thr Asp	
235 240 245	
tat tta tca gga ata tca ctt aca aaa tta aga gat aaa ctc aat gaa	1362
Tyr Leu Ser Gly Ile Ser Leu Thr Lys Leu Arg Asp Lys Leu Asn Glu	
250 255 260 265	
tct ggg cac atc ggt aaa gat ata ccg tgg tct tat cgt acc cta aga	1410
Ser Gly His Ile Gly Lys Asp Ile Pro Trp Ser Tyr Arg Thr Leu Arg	
270 275 280	
caa aca ctt gat aat cca gtt tac tgt ggt tat atc aaa ttt aag gac	1458
Gln Thr Leu Asp Asn Pro Val Tyr Cys Gly Tyr Ile Lys Phe Lys Asp	
285 290 295	
agc cta ttt gaa ggt atg cac aaa cca att atc cct tat gag act tat	1506
Ser Leu Phe Glu Gly Met His Lys Pro Ile Ile Pro Tyr Glu Thr Tyr	
300 305 310	
tta aaa gtt caa aaa gag cta gaa gaa aga caa cag cag act tat gaa	1554
Leu Lys Val Gln Lys Glu Leu Glu Glu Arg Gln Gln Gln Thr Tyr Glu	
315 320 325	
aga aat aac aac cct aga cct ttc caa gct aaa tat atg ctg tca ggg	1602
Arg Asn Asn Asn Pro Arg Pro Phe Gln Ala Lys Tyr Met Leu Ser Gly	
330 335 340 345	

22

```

atg gca agg tgc ggt tac tgt gga gca cct tta aaa att gtt ctt ggc 1650
Met Ala Arg Cys Gly Tyr Cys Gly Ala Pro Leu Lys Ile Val Leu Gly
                               350                               355                               360

cac aaa aga aaa gat gga agc cgc act atg aaa tat cac tgt gca aat 1698
His Lys Arg Lys Asp Gly Ser Arg Thr Met Lys Tyr His Cys Ala Asn
                               365                               370                               375

aga ttt cct cga aaa aca aaa gga att aca gta tat aat gac aat aaa 1746
Arg Phe Pro Arg Lys Thr Lys Gly Ile Thr Val Tyr Asn Asp Asn Lys
                               380                               385                               390

aag tgt gat tca gga act tat gat tta agt aat tta gaa aat act gtt 1794
Lys Cys Asp Ser Gly Thr Tyr Asp Leu Ser Asn Leu Glu Asn Thr Val
                               395                               400                               405

att gac aac ctg att gga ttt caa gaa aat aat gac tcc tta ttg aaa 1842
Ile Asp Asn Leu Ile Gly Phe Gln Glu Asn Asn Asp Ser Leu Leu Lys
410                               415                               420                               425

att atc aat ggc aac aac caa cct att ctt gat act tgc tca ttt aaa 1890
Ile Ile Asn Gly Asn Asn Gln Pro Ile Leu Asp Thr Ser Ser Phe Lys
                               430                               435                               440

aag caa att tca cag atc gat aaa aaa ata caa aag aac tct gat ttg 1938
Lys Gln Ile Ser Gln Ile Asp Lys Lys Ile Gln Lys Asn Ser Asp Leu
                               445                               450                               455

tac cta aat gat ttt atc act atg gat gag ttg aaa gat cgt act gat 1986
Tyr Leu Asn Asp Phe Ile Thr Met Asp Glu Leu Lys Asp Arg Thr Asp
                               460                               465                               470

tcc ctt cag gct gag aaa aag ctg ctt aaa gct aag att agc gaa aat 2034
Ser Leu Gln Ala Glu Lys Lys Leu Leu Lys Ala Lys Ile Ser Glu Asn
475                               480                               485

aaa ttt aat gac tct act gat gtt ttt gag tta gtt aaa act cag ttg 2082
Lys Phe Asn Asp Ser Thr Asp Val Phe Glu Leu Val Lys Thr Gln Leu
490                               495                               500                               505

ggc tca att ccg att aat gaa cta tca tat gat aat aaa aag aaa atc 2130
Gly Ser Ile Pro Ile Asn Glu Leu Ser Tyr Asp Asn Lys Lys Lys Ile
                               510                               515                               520

gtc aac aac ctt gta tca aag gtt gat gtt act gct gat aat gta gat 2178
Val Asn Asn Leu Val Ser Lys Val Asp Val Thr Ala Asp Asn Val Asp
                               525                               530                               535

atc ata ttt aaa ttc caa ctc gct taactgcagc tcgacctgca 2222
Ile Ile Phe Lys Phe Gln Leu Ala
                               540                               545

ggcatgccct gctttaatga gatatgcgag acgcctatga tcgcatgata tttgctttca 2282
attctgttgc gcacgttgta aaaaacctga gcatgtgtag ctcagatcct taccgccggg 2342
ttcggttcat tctaataaat atatcaccgg ttactatcgt atttttatga ataataattct 2402
ccgttcaatt tactgattgt ccaagctt 2430

<210> 46
<211> 545
<212> PRT
<213> Künstliche Sequenz
<223> Beschreibung der künstlichen Sequenz: insert of
        binary vector pCB245-37
<400> 46
Met Ser Ala Ser Ser Leu Phe Asn Leu Pro Leu Ile Arg Leu Arg Ser

```

23

1	5	10	15
Leu Ala Leu Ser Ser Ser Phe Ser Ser Phe Arg Phe Ala His Arg Pro			
	20	25	30
Leu Ser Ser Ile Ser Pro Arg Lys Leu Pro Asn Phe Arg Ala Phe Ser			
	35	40	45
Gly Thr Ala Met Thr Asp Thr Lys Asp Gly Ile Pro Met Ala Lys Lys			
	50	55	60
Ala Ala Ile Tyr Thr Arg Val Ser Thr Thr Asn Gln Ala Glu Glu Gly			
	65	70	75
Phe Ser Ile Asp Glu Gln Ile Asp Arg Leu Thr Lys Tyr Ala Glu Ala			
	85	90	95
Met Gly Trp Gln Val Ser Asp Thr Tyr Thr Asp Ala Gly Phe Ser Gly			
	100	105	110
Ala Lys Leu Glu Arg Pro Ala Met Gln Arg Leu Ile Asn Asp Ile Glu			
	115	120	125
Asn Lys Ala Phe Asp Thr Val Leu Val Tyr Lys Leu Asp Arg Leu Ser			
	130	135	140
Arg Ser Val Arg Asp Thr Leu Tyr Leu Val Lys Asp Val Phe Thr Lys			
	145	150	155
Asn Lys Ile Asp Phe Ile Ser Leu Asn Glu Ser Ile Asp Thr Ser Ser			
	165	170	175
Ala Met Gly Ser Leu Phe Leu Thr Ile Leu Ser Ala Ile Asn Glu Phe			
	180	185	190
Glu Arg Glu Asn Ile Lys Glu Arg Met Thr Met Gly Lys Leu Gly Arg			
	195	200	205
Ala Lys Ser Gly Lys Ser Met Met Trp Thr Lys Thr Ala Phe Gly Tyr			
	210	215	220
Tyr His Asn Arg Lys Thr Gly Ile Leu Glu Ile Val Pro Leu Gln Ala			
	225	230	235
Thr Ile Val Glu Gln Ile Phe Thr Asp Tyr Leu Ser Gly Ile Ser Leu			
	245	250	255
Thr Lys Leu Arg Asp Lys Leu Asn Glu Ser Gly His Ile Gly Lys Asp			
	260	265	270
Ile Pro Trp Ser Tyr Arg Thr Leu Arg Gln Thr Leu Asp Asn Pro Val			
	275	280	285
Tyr Cys Gly Tyr Ile Lys Phe Lys Asp Ser Leu Phe Glu Gly Met His			
	290	295	300
Lys Pro Ile Ile Pro Tyr Glu Thr Tyr Leu Lys Val Gln Lys Glu Leu			
	305	310	315
Glu Glu Arg Gln Gln Gln Thr Tyr Glu Arg Asn Asn Asn Pro Arg Pro			
	325	330	335
Phe Gln Ala Lys Tyr Met Leu Ser Gly Met Ala Arg Cys Gly Tyr Cys			
	340	345	350
Gly Ala Pro Leu Lys Ile Val Leu Gly His Lys Arg Lys Asp Gly Ser			
	355	360	365
Arg Thr Met Lys Tyr His Cys Ala Asn Arg Phe Pro Arg Lys Thr Lys			
	370	375	380
Gly Ile Thr Val Tyr Asn Asp Asn Lys Lys Cys Asp Ser Gly Thr Tyr			
	385	390	395
Asp Leu Ser Asn Leu Glu Asn Thr Val Ile Asp Asn Leu Ile Gly Phe			
	405	410	415
Gln Glu Asn Asn Asp Ser Leu Leu Lys Ile Ile Asn Gly Asn Asn Gln			
	420	425	430
Pro Ile Leu Asp Thr Ser Ser Phe Lys Lys Gln Ile Ser Gln Ile Asp			
	435	440	445
Lys Lys Ile Gln Lys Asn Ser Asp Leu Tyr Leu Asn Asp Phe Ile Thr			
	450	455	460

24

Met Asp Glu Leu Lys Asp Arg Thr Asp Ser Leu Gln Ala Glu Lys Lys
 465 470 475 480
 Leu Leu Lys Ala Lys Ile Ser Glu Asn Lys Phe Asn Asp Ser Thr Asp
 485 490 495
 Val Phe Glu Leu Val Lys Thr Gln Leu Gly Ser Ile Pro Ile Asn Glu
 500 505 510
 Leu Ser Tyr Asp Asn Lys Lys Lys Ile Val Asn Asn Leu Val Ser Lys
 515 520 525
 Val Asp Val Thr Ala Asp Asn Val Asp Ile Ile Phe Lys Phe Gln Leu
 530 535 540
 Ala
 545

<210> 47

<211> 485

<212> PRT

<213> Künstliche Sequenz

<223> Beschreibung der künstlichen Sequenz: insert of
 binary vector pCB245-37

<400> 47

Met Ala Lys Lys Ala Ala Ile Tyr Thr Arg Val Ser Thr Thr Asn Gln
 1 5 10 15
 Ala Glu Glu Gly Phe Ser Ile Asp Glu Gln Ile Asp Arg Leu Thr Lys
 20 25 30
 Tyr Ala Glu Ala Met Gly Trp Gln Val Ser Asp Thr Tyr Thr Asp Ala
 35 40 45
 Gly Phe Ser Gly Ala Lys Leu Glu Arg Pro Ala Met Gln Arg Leu Ile
 50 55 60
 Asn Asp Ile Glu Asn Lys Ala Phe Asp Thr Val Leu Val Tyr Lys Leu
 65 70 75 80
 Asp Arg Leu Ser Arg Ser Val Arg Asp Thr Leu Tyr Leu Val Lys Asp
 85 90 95
 Val Phe Thr Lys Asn Lys Ile Asp Phe Ile Ser Leu Asn Glu Ser Ile
 100 105 110
 Asp Thr Ser Ser Ala Met Gly Ser Leu Phe Leu Thr Ile Leu Ser Ala
 115 120 125
 Ile Asn Glu Phe Glu Arg Glu Asn Ile Lys Glu Arg Met Thr Met Gly
 130 135 140
 Lys Leu Gly Arg Ala Lys Ser Gly Lys Ser Met Met Trp Thr Lys Thr
 145 150 155 160
 Ala Phe Gly Tyr Tyr His Asn Arg Lys Thr Gly Ile Leu Glu Ile Val
 165 170 175
 Pro Leu Gln Ala Thr Ile Val Glu Gln Ile Phe Thr Asp Tyr Leu Ser
 180 185 190
 Gly Ile Ser Leu Thr Lys Leu Arg Asp Lys Leu Asn Glu Ser Gly His
 195 200 205
 Ile Gly Lys Asp Ile Pro Trp Ser Tyr Arg Thr Leu Arg Gln Thr Leu
 210 215 220
 Asp Asn Pro Val Tyr Cys Gly Tyr Ile Lys Phe Lys Asp Ser Leu Phe
 225 230 235 240
 Glu Gly Met His Lys Pro Ile Ile Pro Tyr Glu Thr Tyr Leu Lys Val
 245 250 255
 Gln Lys Glu Leu Glu Glu Arg Gln Gln Gln Thr Tyr Glu Arg Asn Asn
 260 265 270
 Asn Pro Arg Pro Phe Gln Ala Lys Tyr Met Leu Ser Gly Met Ala Arg
 275 280 285
 Cys Gly Tyr Cys Gly Ala Pro Leu Lys Ile Val Leu Gly His Lys Arg

25

290		295		300
Lys Asp Gly Ser Arg Thr Met Lys Tyr His Cys	Ala Asn Arg Phe Pro			
305		310		315
Arg Lys Thr Lys Gly Ile Thr Val Tyr Asn Asp	Asn Lys Lys Cys Asp			
	325		330	335
Ser Gly Thr Tyr Asp Leu Ser Asn Leu Glu Asn Thr Val Ile Asp Asn				
	340		345	350
Leu Ile Gly Phe Gln Glu Asn Asn Asp Ser Leu Leu Lys Ile Ile Asn				
	355		360	365
Gly Asn Asn Gln Pro Ile Leu Asp Thr Ser Ser Phe Lys Lys Gln Ile				
	370		375	380
Ser Gln Ile Asp Lys Lys Ile Gln Lys Asn Ser Asp Leu Tyr Leu Asn				
	385		390	395
Asp Phe Ile Thr Met Asp Glu Leu Lys Asp Arg Thr Asp Ser Leu Gln				
	405		410	415
Ala Glu Lys Lys Leu Leu Lys Ala Lys Ile Ser Glu Asn Lys Phe Asn				
	420		425	430
Asp Ser Thr Asp Val Phe Glu Leu Val Lys Thr Gln Leu Gly Ser Ile				
	435		440	445
Pro Ile Asn Glu Leu Ser Tyr Asp Asn Lys Lys Lys Ile Val Asn Asn				
	450		455	460
Leu Val Ser Lys Val Asp Val Thr Ala Asp Asn Val Asp Ile Ile Phe				
	465		470	475
Lys Phe Gln Leu Ala				480
				485

<210> 48

<211> 2077

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: insert of
vector pCB249-24

<400> 48

```

ggtacgtcga cgaagttcct attccgaagt tcctattctc aagaaagtat aggaacttcg 60
taccgtcgag aaaaaaagaa aggagcaata gcaccctctt gatagaacaa gaaaatgatt 120
attgctcctt tcttttcaaa acctcctata gactaggcca ggaattatct gcagttattt 180
gccaactacc ttagtgatct cgcctttcac gtagtggaca aattcttcca actgatctgc 240
gcgcgaggcc aagcgatctt cttcttgccc aagataagcc tgtctagctt caagtatgac 300
gggctgatac tgggcccggc ggcgctccat tgcccagtcg gcagcgacat ccttcggcgc 360
gattttgccg gttactgcgc tgtaccaaatt gcgggacaac gtaagcacta catttcgctc 420
atcgccagcc cagtcgggcg gcgagttcca tagcgttaag gtttcattta gcgcctcaaa 480
tagatcctgt tcaggaaccg gatcaaagag ttccctccgc gctggaccta ccaaggcaac 540
gctatgttct cttgcttttg tcagcaagat agccagatca atgtcgatcg tggctggctc 600
gaagatacct gcaagaatgt cattgcgctg ccattctcca aattgcagtt cgcgcttagc 660
tggaataacgc cacggaatga tgcgtcgtg cacaacaatg gtgacttcta cagcgcggag 720
aatctcgctc tctccagggg aagccgaagt ttccaaaagg tcgttgatcc gtacctccaa 780
ctcgcttaat tgcgagtttt tatttcgttt atttcaatta aggtaactaa aaaactcctt 840
ttaaggagtt tctgagatcc gtagtgcccc aactggggta acctttgagt tctctcagtt 900
gggggcgtag aattagatct cgactctata tatatttata tattaaatat taattaatat 960
gttagataaaa ttaatattag cgatgggggc ttactaaaga aaaatattta tccacctatc 1020
tctatagtat ataggatcta gttgtaggga gggattcatg acaccacaaa cagccatggg 1080
agcttggtat gaacaagatg gattgcacgc aggttctccg gccgcttggg tggagaggct 1140
attcggctat gactgggcac aacagacaat cggctgctct gatgccgcg tgttccggct 1200
gtcagcgcag gggcgcccgg ttctttttgt caagaccgac ctgtccgggt ccctgaatga 1260
actgcaggac gaggcagcgc ggctatcggt gctggccacg acgggcgttc cttgcgcagc 1320

```


26

```

tgtgctcgac gttgtcactg aagcgggaag ggactggctg ctattgggcg aagtgccggg 1380
gcaggatctc ctgtcatctc accttgtctc tgccgagaaa gtatccatca tggctgatgc 1440
aatgcgggcg ctgcatacgc ttgatccggc tacctgcccc ttccgaccacc aagcgaaaca 1500
tcgcatcgag cgagcacgta ctcgatgga agccggtctt gtcgatcagg atgatctgga 1560
cgaagagcat caggggctcg cgccagccga actgttcgcc aggtcaagg cgcgcagtc 1620
cgacggcgag gatctcgctg tgacctatgg cgatgcctgc ttgccgaata tcatggtgga 1680
aaatggccgc ttttctggat tcatcgactg tggccggctg ggtgtggcg accgctatca 1740
ggacatagcg ttggctaccc gtgatattgc tgaagagctt ggcggcgaat gggctgaccg 1800
cttctcgtg ctttacggta tcgccctcc cgattcgag cgcctgcct tctatgcct 1860
tcttgacgag ttctcttgag cgggacccaa gctagcttcg acggatcaaa gaattcagta 1920
gacattagca gataaattag caggaaataa agaaggataa ggagaaagaa ctcaagtaat 1980
tatccttcgt tctcttaatt gaattgcaat taaactcggc ccaatctttt actaaaagga 2040
ttgagccgaa tactcgagag ggtacgtac cgagctc 2077

```

<210> 49

<211> 2662

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: insert of vector pCB384-40

<400> 49

```

ggtagctcga cgaagttcct attccgaagt tcctattctc aagaaagtat aggaacttcg 60
taccgtcgag aaaaaaagaa aggagcaata gcacctctt gatagaacaa gaaaatgatt 120
attgctcctt tcttttcaaa acctcctata gactaggcca ggaattatct gcagttatct 180
gcccaactacc ttagtgatct cgctttcac gtagtggaca aattcttcca actgatctgc 240
gcgcgaggcc aagcgatctt cttctgtcc aagataagcc tgtctagctt caagtatgac 300
gggctgatac tgggcccggc ggcgtccat tgcccagtcg gcagcgacat ccttcggcgc 360
gattttgccg gttactgcgc tgtacaaat gcgggacaac gtaagcacta catttcgctc 420
atcgccagcc cagtcgggag gcgagttcca tagcgttaag gtttcattta gcgcctcaaa 480
tagatcctgt tcaggaaccg gatcaaagag ttctccgcc gctggacctt ccaaggcaac 540
gctatgttct cttgcttttg tcagcaagat agccagatca atgtcgatcg tggctggctc 600
gaagatacct gcaagaatgt cattgcgtg ccattctcca aattgcagtt cgcgcttagc 660
tggtataacgc cacggaatga tgcgtcgtg cacaacaatg gtgacttcta cagcgccgag 720
aatctcgtc tctccagggg aagccgaagt ttccaaaagg tcgttgatcc gtacctccaa 780
ctcgcttaat tgcgagtttt tatttcgttt atttcaatta aggttaactaa aaaactcctt 840
ttaaggagtt tctgagatcc gtagtgccc aactgggta accttgagt tctctcagtt 900
gggggcgtag aattagatct cgactctata tatatttata tattaatat taattaatat 960
gttagataaaa ttaatatag cgatggggtc ttactaaaga aaaatattta tccacctatc 1020
tctatagtat ataggatcta gttgtaggga gggattcatg acaccacaaa cagccatggg 1080
agcttggatt gaacaagatg gattgcagc aggttctccg gccgcttggg tggagaggct 1140
attcggtat gactgggcac aacagacaat cggctgctct gatgccgccc tgttccggct 1200
gtcagcgag gggcgcccgc ttctttttgt caagaccgac ctgtccggtg cctgaatga 1260
actgcaggac gaggcagcgc ggctatcgtg gctggccacg acggcggttc cttgcgcagc 1320
tgtgctcgac gttgtcactg aagcgggaag ggactggctg ctattgggcg aagtgccggg 1380
gcaggatctc ctgtcatctc accttgtctc tgccgagaaa gtatccatca tggctgatgc 1440
aatgcgggcg ctgcatacgc ttgatccggc tacctgcccc ttccgaccacc aagcgaaaca 1500
tcgcatcgag cgagcacgta ctcgatgga agccggtctt gtcgatcagg atgatctgga 1560
cgaagagcat caggggctcg cgccagccga actgttcgcc aggtcaagg cgcgcagtc 1620
cgacggcgag gatctcgctg tgacctatgg cgatgcctgc ttgccgaata tcatggtgga 1680
aaatggccgc ttttctggat tcatcgactg tggccggctg ggtgtggcg accgctatca 1740
ggacatagcg ttggctaccc gtgatattgc tgaagagctt ggcggcgaat gggctgaccg 1800
cttctcgtg ctttacggta tcgccctcc cgattcgag cgcctgcct tctatgcct 1860
tcttgacgag ttctcttgag cgggacggat cctgatcttc cattttctat tttgatttgt 1920
agaaaactag tgtgcttggg agtccctgat gattaaataa accaagattt taccatggcg 1980
ctcaccaatg ctcaaatctt ggctgtgatt gacagttggg aagaaacagt cggctcagttt 2040
ccagtataa cgcaccatgt accattaggt ggcggtctgc aaggaaacgct ccattgttac 2100

```

27

```

gagatccccc tagcagctcc ttatgggggtt ggcttttgcta agaatggggcc taccgcgtgg 2160
caatacaaac ggacaatcaa tcaagtcgtc cacagatggg ggtcccacac agtcccctttt 2220
ctattagaac cggataacat caacggcaaa acctgcacag catcgcacct atgtcataat 2280
actcgatgcc acaatccctt gcaattgtgc tgggagtcac tagacgacaa caaaggcaga 2340
aactgggtgcc ccggtcccaa cgggggatgt gtccatgcgg tggtttggtt aaggcagggt 2400
ccgtttgtacg gccagggggc gactgtggca ggtcctcaac aaagggggcag tcactttgtg 2460
gtataactgc agaagcttca attgcatgct ctagagaatt cagtagacat tagcagataa 2520
attagcagga aataaagaag gataaggaga aagaactcaa gtaattatcc ttcgtttctt 2580
taattgaatt gcaattaaac tcggcccaat cttttactaa aaggattgag ccgaatactc 2640
gagaggggta cgtaccgagc tc 2662

```

<210> 50

<211> 1287

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: insertion of
vector pCB456-2

<400> 50

```

gagctctcag tactcgagac atttcaccct agaaatagac ttaaacttta ctggcttaac 60
tttaaccttg gaccataaaa ggcaccccat tgcgagagtg cccttgatta accaaatgaa 120
acgaagtcta accgaaagca gttatggcaa tggcagaatt ctgatcacgg aagatagctt 180
tggcaaaaaa agcaaaaagc atttaccttg attgagatgt taattgtgtt ggcaattatc 240
agtattttta ttttgctttt tgtgccaaat ttgatactag agcttcgggt gccagggcgt 300
gcccttgggc tccccgggcg cgtactcgac gctaccttaa gagagtcaag ctaattctaa 360
tctgcagtct agcgtgcggc cgctctagaa ctagtggatc ccccgggctg cagttatttg 420
ccaactacct tagtgatctc gcctttcacg tagtggacaa attcttccaa ctgatctgcg 480
cgcgaggcca agcgatcttc ttcttgctca agataagcct gtctagcttc aagtatgacg 540
ggctgatact gggccggcag gcgctccatt gccagtcgg cagcgacatc cttcggcgcg 600
atcttgccgg ttactgcgct gtaccaaagc cgggacaacg taagcactac atttcgctca 660
tcgccagccc agtcggggcg cgagttccat agcggttaagg tttcatttag cgcctcaaat 720
agatcctggt caggaaccgg atcaaagagt tcctccgccc ctggacctac caaggcaacg 780
ctatgttctc ttgcttttgt cagcaagata gccagatcaa tgtcgatcgt ggctggctcg 840
aagatacctg caagaatgac attgcgctgc cattctccaa attgcagttc gcgcttagct 900
ggataacgcc acggaatgat gtcgtcgtgc acaacaatgg tgacttctac agcgcgagga 960
atctcgtctc ctccagggga agccgaagtt tccaaaaggt cgttgatcaa agctcgccgc 1020
gttggtttcat caagccttac ggtaaccgta accagcaaat caatatcact gtgtggcttc 1080
aggccgccat ccaactgcga gccgtacaaa tgtacggcca gcaacgctcg ttcgagatgg 1140
cgctcgatga cgccaactac ctctgatagt tgagttgata cttcggcgat aaccgcttca 1200
cgagccatga gatcctccag atccatgtat cattatagat aattgaagag tgaatgtcaa 1260
gtcgacctcg aggggggggc cgggtacc 1287

```

<210> 51

<211> 45

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:
oligonucleotide primer

<400> 51

```

cggatccgga ggcacatga ctaagaaagt agcaatctat acacg

```

45

<210> 52

<211> 22

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:
oligonucleotide primer

<400> 52

ataagcttta ttctcgatat cg

22

<210> 53

<211> 1845

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: insertion of
vector pCB557-1

<400> 53

```

ggtacctcca actcgcttaa ttgcgagttt ttatttcggt tatttcaatt aaggtaacta 60
aaaaactcct ttttaaggagt ttctgagatc aattctacgc cccaactga gagaactcaa 120
aggttacccc agttggggca ctacggatca gagatgtaac tccagttcct tcggaatcgg 180
tagtcaatcc tatttccgat aggggcagtt gacaattgaa tccgatittg accattattt 240
tcatatccgt aatagtgcga aaagaaggcc cggctccaag ttgttcaaga atagtggcgt 300
tgagttttct gaccctttga cttaggatta gtcagttcta ttctcgatg gggcggggaa 360
gggatataac tcagcggtag agtgtcacct tgacgtgggt gaagtcacat gttcgagcct 420
gattatccct aagcccaatg tgagtttttc tagttggatt ttgacaatt ttcctctgaa 480
ttatataatt aacataggat caattcatct cgaggaggtc acatgggagc ttggattgaa 540
caagatggat tgcacgcagg ttctccggcc gcttgggtgg agaggctatt cggctatgac 600
tgggcacaac agacaatcgg ctgctctgat gccgcggtgt tccggctgtc agcgcagggg 660
cgccccgttc tttttgtcaa gaccgacctg tccggtgccc tgaatgaact gcaggacgag 720
gcagcgcggc tatcgtggct ggccacgacg ggcggttcct gcgcagctgt gctcgacgtt 780
gtcactgaag cgggaaggga ctggctgcta ttgggcgaag tgccggggca ggatctcctg 840
tcatctcacc ttgctcctgc cgagaaagta tccatcatgg ctgatgcaat gcggcggtctg 900
catacgtttg atccggctac ctgcccattc gaccaccaag cgaaacatcg catcgagcga 960
gcacgtactc ggatggaagc cggctctgtc gatcaggatg atctggacga agagcatcag 1020
gggctcgcgc cagccgaact gttcgccagg ctcaaggcgc gcatgcccga cggcgaggat 1080
ctcgtcgtga cccatggcga tgccctgctt ccgaatatca tggtggaana tggccgcttt 1140
tctggattca tgcactgtgg ccggctgggt gtggcggacc gctatcagga catagcggtt 1200
gctacccgtg atattgctga agagcttggc gggaatggg ctgaccgctt cctcgtgctt 1260
tacggtatcg ccgctcccga ttgcagcgc atcgcttct atcgcttct tgacgagttc 1320
ttctgatcta gaggaggtca catggcgctc accaatgctc aaatcttggc tgtgattgac 1380
agttgggaag aaacagtcgg tcagtttcca gtgataacgc accatgtacc attaggtggc 1440
ggtctgcaag gaacgctcca ttgttacgag atccccctag cagctcctta tgggggttggc 1500
tttgctaaga atgggcctac ccgctggcaa taaaaacgga caatcaatca agtcgtccac 1560
agatgggggt cccacacagt cccttttcta ttagaaccgg ataacatcaa cggcaaaacc 1620
tgcacagcat cgcacctatg tcataatact cgatgccaca atcccttgca cttgtgctgg 1680
gagtcactag acgacaacaa aggcagaaac tggtgccccg gtcccaacgg gggatgtgtc 1740
catgcggtgg tttgtttaag gcagggtccg ttgtacggcc caggggagac tgtggcaggt 1800
cctcaacaaa ggggcagtcg ctttgtggta taaggtagcg agtc 1845

```

<210> 54

<211> 1965

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: insertion of
vector pCB554-3

<400> 54

```

ggtaccgggc cccccctoga ggtcgacttg acattcaact ttcaattatc tataatgata 60
catggatccg gaggcacatc gacacaagggt gttgtgaccg ggggtggacac gtacgggggt 120
gcttacgacc gtcagtcgcg cgagcgcgag aattcgagcg cagcaagccc agcgacacag 180

```

```

cgtagcgcca acgaagacaa ggcggccgac cttcagcgcg aagtcgagcg cgacggggggc 240
cggttcaggt tcgtcgggca tttcagcgaa gcgccgggca cgtcggcggt cgggacggcg 300
gagcgccccg agttcgaacg catcctgaac gaatgccgcg ccgggcggtt caacatgac 360
attgtctatg acgtgtcgcg cttctcgcgc ctgaagggtca tggacgcgat tccgattgtc 420
tcggaattgc tcgccctggg cgtgacgatt gtttccactc aggaaggcgt cttccggcag 480
ggaaacgtca tggacctgat tcacctgatt atgcggctcg acgcgtcgca caaagaatct 540
tcgctgaagt cggcgaagat tctcgacacg aagaaccttc agcgcgaatt gggcgggtac 600
gtcggcgggg aggcgcctta cggcttcgag cttgtttcgg agacgaagga gatcacgcgc 660
aacggccgaa tggatcaatgt cgtcatcaac aagcttgccg actcgaccac tccccttacc 720
ggacccttcg agttcgagcc cgacgtaatc cgggtggtggt ggcgtgagat caagacgcac 780
aaacaccttc cttcaagcc gggcagtcga gccgccattc acccgggcag catcacgggg 840
ctttgtaagc gcatggacgc tgacgccgtg ccgaccggg gcgagacgat tgggaagaag 900
accgcttcaa gcgcctggga cccggcaacc gttatgcgaa tccttcggga cccgcgtatt 960
gcgggcttcg ccgctgaggt gatctacaag aagaagccgg acggcacgcc gaccacgaag 1020
attgaggggt accgcattca gcgcgaccgg atcacgctcc ggccgggtcg gcttgattgc 1080
ggaccgatca tcgagcccg ctagtggtat gagcttcagg cgtggttggg cggcaggggg 1140
cgccgcaagg ggttttcccg ggggcaagcc attctgtccg ccatggacaa gctgtactgc 1200
gagtgtggcg ccgtcatgac ttcgaagcgc ggggaagaat cgatcaagga ctcttaccgc 1260
tgccgtcgcc ggaagtggt cgaccgtcc gcacctgggc agcacgaagg cacgtgcaac 1320
gtcagcatgg cggcactcga caagtgcgtt gcggaacgca tcttcaacaa gatcaggcac 1380
gccgaaggcg acgaagagac gttggcgctt ctgtgggaag ccgcccgacg cttcggcaag 1440
ctcactgagg cgctgagaa gagcggcgaa cgggcgaacc ttgttgcgga gcgcgccgac 1500
gccctgaacg cccttgaaga gctgtacgaa gaccgcgcgg caggcgcgta cgacggaccc 1560
gttggcagga agcacttccg gaagcaacag gcagcgctga cgctccggca gcaaggggag 1620
gaagagcggc ttgccgaact tgaagccgcc gaagccccga agcttccct tgaccaatgg 1680
ttccccgaag acgccgacgc tgacccgacc ggccctaagt cgtggtgggg gcgcgcgtca 1740
gtagacgaca agcgcgtgtt cgtcgggctc ttcgtagaca agatcggtgt cacgaagtcg 1800
actacgggca gggggcaggg aacgcccac gagaaagcgc cttcgatcac gtgggcgaag 1860
ccgccgaccg acgacgacga agacgacgcc caggacggca cgggaagacgt agcggcgtag 1920
ctgcagctcg agatctctag agcggccgcc accgcggtgg agtc 1965

```

<210> 55

<211> 42

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:
oligonucleotide primer

<400> 55

tcgacttgac attcactctt caattatcta taatgatata tg

42

<210> 56

<211> 26

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:
oligonucleotide primer

<400> 56

cctcgagctg cagttaagcg agttgg

26

<210> 57

<211> 1527

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: PCR produkt
TP-Rek

<400> 57

```

tcgacttgac attcactctt caattatcta taatgatata tggatccgga ggcacatga 60
ctaagaaagt agcaatctat acacgagtat ccactactaa ccaagcagag gaaggggttct 120
caattgatga gcaaattgac cgtttaacaa aatatgctga agcaatgggg tggcaagtat 180
ctgatactta tactgatgct gggttttcag gggccaaact tgaacgcca gcaatgcaaa 240
gattaatcaa cgatatcgag aataaagctt ttgatacagt tcttgatat aagctagacc 300
gcctttcacg tagtgtaaga gatactcttt atcttggtta ggatgtgttc acaaaaaata 360
aaatagactt tatctcgctt aatgaaagta ttgatacttc ttctgctatg ggtagcttgt 420
ttctcactat tctttctgca attaatgagt ttgaaagaga gaatataaaa gaacgcagta 480
ctatgggttaa actagggcga gcgaaatctg gtaagtctat gatgtggact aagacagctt 540
ttgggtatta ccacaacaga aagacaggta tattagaaat tggttcctta caagctacaa 600
tagttgaaca aatattcact gattatttat caggaatatc acttacaata ttaagagata 660
aactcaatga atctgggcac atcggtaaa atataccgtg gtcttatcgt accctaagac 720
aaacacttga taatccagtt tactgtggtt atatcaaatt taaggacagc ctatttgaag 780
gtagtcacaa accaattatc cttatgaga cttattttaa agttcaaaaa gagctagaag 840
aaagacaaca gcagacttat gaaagaaata acaaccctag acctttccaa gctaaatata 900
tgctgtcagg gatggcaagg tgcggttact gtggagcacc tttaaaaatt gttcttggcc 960
acaaaaagaaa agatggaagc cgcactatga aatatcactg tgcaaataga tttcctcgaa 1020
aaacaaaagg aattacagta tataatgaca ataaaaagtg tgattcagga acttatgatt 1080
taagtaattt agaaaatact gttattgaca acctgattgg atttcaagaa aataatgact 1140
ccttattgaa aattatcaat ggcaacaacc aacctattct tgatacttcg tcatttataa 1200
agcaaatctt acagatcgat aaaaaatac aaaagaactc tgatttgtac ctaaatgatt 1260
ttatcactat ggatgagttg aaagatcgta ctgattccct tcaggctgag aaaaagctgc 1320
ttaagctaa gattagcgaa aataaattta atgactctac tgatgtttt gagttagtta 1380
aaactcagtt gggctcaatt ccgattaatg aactatcata tgataataa aagaaaatcg 1440
tcaacaacct tgtatcaaag gttgatgta ctgctgataa tgtagatatc atatttaaat 1500
tccaactcgc ttaactgcag ctcgagg 1527

```

<210> 58

<211> 1213

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: probe for southern blot

<400> 58

```

ttatcacagg cagtctgttc aggggttccaa actcaacgat ggcaactaaa cagcaggggtt 60
gcgctcggtg cgggacttaa cccaacacct tacggcacga gctgacgaca gccatgcacc 120
acctgtgtcc gcgttcccga aggcaccctt ctctttcaag aggattcgcg gcatgtcaag 180
ccctggtaag gttcttcgct ttgcatcgaa taaaccaca tgctccaccg cttgtgcggg 240
ccccgctcaa ttccttttag tttcattctt gcgaacgtac tccccaggcg ggatacttaa 300
cgcgttagct acagcactgc acgggtcgat acgcacagcg cctagtatcc atcgtttacg 360
gctaggacta ctggggatc taatcccatt cgctccccta gctttcgtct ctcaagtgtca 420
gtgtcggccc agcagagtgc tttcgccgtt ggtgttcttt ccgatctcta cgcatttcac 480
cgctccaccg gaaattccct ctgcccctac cgtactccag cttggtagtt tccaccgcct 540
gtccagggtt gagccctggg atttgacggc ggacttaaaa agccacctac agacgcttta 600
cgcccaatca ttcgggataa cgcttgcatc ctctgtatta ccgcggtgctc tggcagagag 660
ttagccgatg cttattcccc agataccgtc attgcttctt ctccgggaaa agaagttcac 720
gaccggtggg ccttctacct ccacgcggca ttgctcgtc aggccttcgc ccattgcgga 780
aaattcccca ctgctgcctc ccgtaggagt ctgggcccgtg tctcagtcct agtgtggctg 840
atcatcctct cggaccagct actgatcatc gccttggtta gctattgcct caccaactag 900
ctaactcagc cgtttccagc tgttggtccc ctcccaaggg caggttctta cgcgttactc 1020
accggtccgc cactggaaac accacttccc gtcogacttg catgtgttaa gcatgccgcc 1080
agcgttcata ctgagccagg atcgaactct coatgagatt catagttgca ttacttatag 1140
cttcttgggt cgtagacaaa gcggattcgg aattgtcttt cattccaagg cataacttgt 1200
atccatgcgc ttc 1213

```

<210> 59
 <211> 34
 <212> DNA
 <213> Künstliche Sequenz
 <220>
 <223> Beschreibung der künstlichen Sequenz: : promoter
 sequence derived from consensus sequence of
 E.coli sigma70 promoters
 <400> 59
 ttgacaattt tcctctgaat tatataatta acat 34
 <210> 60
 <211> 1564
 <212> DNA
 <213> Künstliche Sequenz
 <220>
 <223> Beschreibung der künstlichen Sequenz: Insertion of
 vector pCB487-20
 <220>
 <221> misc_feature
 <222> (28)..(32)
 <223> ribosome binding site
 <220>
 <221> CDS
 <222> (37)..(1491)
 <223> coding for TP901 rekombinase
 <400> 60
 gcggccgcga attcgccctt cggatccgga ggcattc atg act aag aaa gta gca 54
 Met Thr Lys Lys Val Ala
 1 5
 atc tat aca cga gta tcc act act aac caa gca gag gaa ggg ttc tca 102
 Ile Tyr Thr Arg Val Ser Thr Thr Asn Gln Ala Glu Glu Gly Phe Ser
 10 15 20
 att gat gag caa att gac cgt tta aca aaa tat gct gaa gca atg ggg 150
 Ile Asp Glu Gln Ile Asp Arg Leu Thr Lys Tyr Ala Glu Ala Met Gly
 25 30 35
 tgg caa gta tct gat act tat act gat gct ggt ttt tca ggg gcc aaa 198
 Trp Gln Val Ser Asp Thr Tyr Thr Asp Ala Gly Phe Ser Gly Ala Lys
 40 45 50
 ctt gaa cgc cca gca atg caa aga tta atc aac gat atc gag aat aaa 246
 Leu Glu Arg Pro Ala Met Gln Arg Leu Ile Asn Asp Ile Glu Asn Lys
 55 60 65 70
 gct ttt gat aca gtt ctt gta tat aag cta gac cgc ctt tca cgt agt 294
 Ala Phe Asp Thr Val Leu Val Tyr Lys Leu Asp Arg Leu Ser Arg Ser
 75 80 85
 gta aga gat act ctt tat ctt gtt aag gat gtg ttc aca aaa aat aaa 342
 Val Arg Asp Thr Leu Tyr Leu Val Lys Asp Val Phe Thr Lys Asn Lys
 90 95 100
 ata gac ttt atc tcg ctt aat gaa agt att gat act tct tct gct atg 390
 Ile Asp Phe Ile Ser Leu Asn Glu Ser Ile Asp Thr Ser Ser Ala Met
 105 110 115
 ggt agc ttg ttt ctc act att ctt tct gca att aat gag ttt gaa aga 438
 Gly Ser Leu Phe Leu Thr Ile Leu Ser Ala Ile Asn Glu Phe Glu Arg
 120 125 130

32

gag aat ata aaa gaa cgc atg act atg ggt aaa cta ggg cga gcg aaa	486
Glu Asn Ile Lys Glu Arg Met Thr Met Gly Lys Leu Gly Arg Ala Lys	
135 140 145 150	
tct ggt aag tct atg atg tgg act aag aca gct ttt ggg tat tac cac	534
Ser Gly Lys Ser Met Met Trp Thr Lys Thr Ala Phe Gly Tyr Tyr His	
155 160 165	
aac aga aag aca ggt ata tta gaa att gtt cct tta caa gct aca ata	582
Asn Arg Lys Thr Gly Ile Leu Glu Ile Val Pro Leu Gln Ala Thr Ile	
170 175 180	
gtt gaa caa ata ttc act gat tat tta tca gga ata tca ctt aca aaa	630
Val Glu Gln Ile Phe Thr Asp Tyr Leu Ser Gly Ile Ser Leu Thr Lys	
185 190 195	
tta aga gat aaa ctc aat gaa tct ggg cac atc ggt aaa gat ata ccg	678
Leu Arg Asp Lys Leu Asn Glu Ser Gly His Ile Gly Lys Asp Ile Pro	
200 205 210	
tgg tct tat cgt acc cta aga caa aca ctt gat aat cca gtt tac tgt	726
Trp Ser Tyr Arg Thr Leu Arg Gln Thr Leu Asp Asn Pro Val Tyr Cys	
215 220 225 230	
ggg tat atc aaa ttt aag gac agc cta ttt gaa ggt atg cac aaa cca	774
Gly Tyr Ile Lys Phe Lys Asp Ser Leu Phe Glu Gly Met His Lys Pro	
235 240 245	
att atc cct tat gag act tat tta aaa gtt caa aaa gag cta gaa gaa	822
Ile Ile Pro Tyr Glu Thr Tyr Leu Lys Val Gln Lys Glu Leu Glu Glu	
250 255 260	
aga caa cag cag act tat gaa aga aat aac aac cct aga cct ttc caa	870
Arg Gln Gln Gln Thr Tyr Glu Arg Asn Asn Asn Pro Arg Pro Phe Gln	
265 270 275	
gct aaa tat atg ctg tca ggg atg gca agg tgc ggt tac tgt gga gca	918
Ala Lys Tyr Met Leu Ser Gly Met Ala Arg Cys Gly Tyr Cys Gly Ala	
280 285 290	
cct tta aaa att gtt ctt ggc cac aaa aga aaa gat gga agc cgc act	966
Pro Leu Lys Ile Val Leu Gly His Lys Arg Lys Asp Gly Ser Arg Thr	
295 300 305 310	
atg aaa tat cac tgt gca aat aga ttt cct cga aaa aca aaa gga att	1014
Met Lys Tyr His Cys Ala Asn Arg Phe Pro Arg Lys Thr Lys Gly Ile	
315 320 325	
aca gta tat aat gac aat aaa aag tgt gat tca gga act tat gat tta	1062
Thr Val Tyr Asn Asp Asn Lys Lys Cys Asp Ser Gly Thr Tyr Asp Leu	
330 335 340	
agt aat tta gaa aat act gtt att gac aac ctg att gga ttt caa gaa	1110
Ser Asn Leu Glu Asn Thr Val Ile Asp Asn Leu Ile Gly Phe Gln Glu	
345 350 355	
aat aat gac tcc tta ttg aaa att atc aat ggc aac aac caa cct att	1158
Asn Asn Asp Ser Leu Leu Lys Ile Ile Asn Gly Asn Asn Gln Pro Ile	
360 365 370	
ctt gat act tcg tca ttt aaa aag caa att tca cag atc gat aaa aaa	1206
Leu Asp Thr Ser Ser Phe Lys Lys Gln Ile Ser Gln Ile Asp Lys Lys	
375 380 385 390	
ata caa aag aac tct gat ttg tac cta aat gat ttt atc act atg gat	1254
Ile Gln Lys Asn Ser Asp Leu Tyr Leu Asn Asp Phe Ile Thr Met Asp	
395 400 405	

33

gag ttg aaa gat cgt act gat tcc ctt cag gct gag aaa aag ctg ctt 1302
 Glu Leu Lys Asp Arg Thr Asp Ser Leu Gln Ala Glu Lys Lys Leu Leu
 410 415 420

aaa gct aag att agc gaa aat aaa ttt aat gac tct act gat gtt ttt 1350
 Lys Ala Lys Ile Ser Glu Asn Lys Phe Asn Asp Ser Thr Asp Val Phe
 425 430 435

gag tta gtt aaa act cag ttg ggc tca att ccg att aat gaa cta tca 1398
 Glu Leu Val Lys Thr Gln Leu Gly Ser Ile Pro Ile Asn Glu Leu Ser
 440 445 450

tat gat aat aaa aag aaa atc gtc aac aac ctt gta tca aag gtt gat 1446
 Tyr Asp Asn Lys Lys Lys Ile Val Asn Asn Leu Val Ser Lys Val Asp
 455 460 465 470

gtt act gct gat aat gta gat atc ata ttt aaa ttc caa ctc gct 1491
 Val Thr Ala Asp Asn Val Asp Ile Ile Phe Lys Phe Gln Leu Ala
 475 480 485

taactgcagc tcgaggaagg gcgaattctg cagatatecca tcacactggc ggccgctoga 1551
 gcatgcatct agt 1564

<210> 61
 <211> 485
 <212> PRT
 <213> Künstliche Sequenz
 <223> Beschreibung der künstlichen Sequenz: Insertion of
 vector pCB487-20

<400> 61
 Met Thr Lys Lys Val Ala Ile Tyr Thr Arg Val Ser Thr Thr Asn Gln
 1 5 10 15
 Ala Glu Glu Gly Phe Ser Ile Asp Glu Gln Ile Asp Arg Leu Thr Lys
 20 25 30
 Tyr Ala Glu Ala Met Gly Trp Gln Val Ser Asp Thr Tyr Thr Asp Ala
 35 40 45
 Gly Phe Ser Gly Ala Lys Leu Glu Arg Pro Ala Met Gln Arg Leu Ile
 50 55 60
 Asn Asp Ile Glu Asn Lys Ala Phe Asp Thr Val Leu Val Tyr Lys Leu
 65 70 75 80
 Asp Arg Leu Ser Arg Ser Val Arg Asp Thr Leu Tyr Leu Val Lys Asp
 85 90 95
 Val Phe Thr Lys Asn Lys Ile Asp Phe Ile Ser Leu Asn Glu Ser Ile
 100 105 110
 Asp Thr Ser Ser Ala Met Gly Ser Leu Phe Leu Thr Ile Leu Ser Ala
 115 120 125
 Ile Asn Glu Phe Glu Arg Glu Asn Ile Lys Glu Arg Met Thr Met Gly
 130 135 140
 Lys Leu Gly Arg Ala Lys Ser Gly Lys Ser Met Met Trp Thr Lys Thr
 145 150 155 160
 Ala Phe Gly Tyr Tyr His Asn Arg Lys Thr Gly Ile Leu Glu Ile Val
 165 170 175
 Pro Leu Gln Ala Thr Ile Val Glu Gln Ile Phe Thr Asp Tyr Leu Ser
 180 185 190
 Gly Ile Ser Leu Thr Lys Leu Arg Asp Lys Leu Asn Glu Ser Gly His
 195 200 205

34

Ile	Gly	Lys	Asp	Ile	Pro	Trp	Ser	Tyr	Arg	Thr	Leu	Arg	Gln	Thr	Leu	210	215	220	
Asp	Asn	Pro	Val	Tyr	Cys	Gly	Tyr	Ile	Lys	Phe	Lys	Asp	Ser	Leu	Phe	225	230	235	240
Glu	Gly	Met	His	Lys	Pro	Ile	Ile	Pro	Tyr	Glu	Thr	Tyr	Leu	Lys	Val	245	250	255	
Gln	Lys	Glu	Leu	Glu	Glu	Arg	Gln	Gln	Gln	Thr	Tyr	Glu	Arg	Asn	Asn	260	265	270	
Asn	Pro	Arg	Pro	Phe	Gln	Ala	Lys	Tyr	Met	Leu	Ser	Gly	Met	Ala	Arg	275	280	285	
Cys	Gly	Tyr	Cys	Gly	Ala	Pro	Leu	Lys	Ile	Val	Leu	Gly	His	Lys	Arg	290	295	300	
Lys	Asp	Gly	Ser	Arg	Thr	Met	Lys	Tyr	His	Cys	Ala	Asn	Arg	Phe	Pro	305	310	315	320
Arg	Lys	Thr	Lys	Gly	Ile	Thr	Val	Tyr	Asn	Asp	Asn	Lys	Lys	Cys	Asp	325	330	335	
Ser	Gly	Thr	Tyr	Asp	Leu	Ser	Asn	Leu	Glu	Asn	Thr	Val	Ile	Asp	Asn	340	345	350	
Leu	Ile	Gly	Phe	Gln	Glu	Asn	Asn	Asp	Ser	Leu	Leu	Lys	Ile	Ile	Asn	355	360	365	
Gly	Asn	Asn	Gln	Pro	Ile	Leu	Asp	Thr	Ser	Ser	Phe	Lys	Lys	Gln	Ile	370	375	380	
Ser	Gln	Ile	Asp	Lys	Lys	Ile	Gln	Lys	Asn	Ser	Asp	Leu	Tyr	Leu	Asn	385	390	395	400
Asp	Phe	Ile	Thr	Met	Asp	Glu	Leu	Lys	Asp	Arg	Thr	Asp	Ser	Leu	Gln	405	410	415	
Ala	Glu	Lys	Lys	Leu	Leu	Lys	Ala	Lys	Ile	Ser	Glu	Asn	Lys	Phe	Asn	420	425	430	
Asp	Ser	Thr	Asp	Val	Phe	Glu	Leu	Val	Lys	Thr	Gln	Leu	Gly	Ser	Ile	435	440	445	
Pro	Ile	Asn	Glu	Leu	Ser	Tyr	Asp	Asn	Lys	Lys	Lys	Ile	Val	Asn	Asn	450	455	460	
Leu	Val	Ser	Lys	Val	Asp	Val	Thr	Ala	Asp	Asn	Val	Asp	Ile	Ile	Phe	465	470	475	480
Lys	Phe	Gln	Leu	Ala												485			

INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP 02/14303

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N15/82 C12N5/14

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, BIOSIS, WPI Data, PAJ, EMBASE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 01 29241 A (CALGENE LLC) 26 April 2001 (2001-04-26)	1-5,10, 12-14
Y	siehe das ganze Dokument	7-9,11
X	WO 01 07572 A (UNIV CALIFORNIA ;US AGRICULTURE (US)) 1 February 2001 (2001-02-01) siehe insbes. Seiten 12-16	7,9, 11-14
Y	WO 00 11155 A (UNIV LELAND STANFORD JUNIOR) 2 March 2000 (2000-03-02) siehe S. 5, letzter Abschnitt, S. 30-36, Ansprüche	7-9,11
	--- -/--	

☒ Further documents are listed in the continuation of box C.☒ Patent family members are listed in annex.

* Special categories of cited documents:

A document defining the general state of the art which is not considered to be of particular relevance

E earlier document but published on or after the international filing date

L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

O document referring to an oral disclosure, use, exhibition or other means

P document published prior to the international filing date but later than the priority date claimed

T later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

X document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

Y document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

Z document member of the same patent family

Date of the actual completion of the international search

31 March 2003

Date of mailing of the international search report

11/04/2003

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Grosskopf, R

INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP 02/14303

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CORNEILLE SYLVIE ET AL: "Efficient elimination of selectable marker genes from the plastid genome by the CRE-lox site-specific recombination system." PLANT JOURNAL, vol. 27, no. 2, July 2001 (2001-07), pages 171-178, XP002236641 ISSN: 0960-7412 siehe S. 172, linke Kolonne ---	7-10
A	BOGORAD L: "Engineering chloroplasts: an alternative site for foreign genes, proteins, reactions and products" TRENDS IN BIOTECHNOLOGY, ELSEVIER, AMSTERDAM, NL, vol. 18, no. 6, June 2000 (2000-06), pages 257-263, XP004203651 ISSN: 0167-7799 -----	

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/EP 02/14303

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
WO 0129241	A	26-04-2001	AU 1090701 A BR 0014858 A EP 1222296 A2 WO 0129241 A2	30-04-2001 16-07-2002 17-07-2002 26-04-2001
WO 0107572	A	01-02-2001	AU 6364800 A EP 1234022 A2 JP 2003505065 T WO 0107572 A2	13-02-2001 28-08-2002 12-02-2003 01-02-2001
WO 0011155	A	02-03-2000	AU 5898599 A WO 0011155 A1	14-03-2000 02-03-2000

INTERNATIONALER RECHERCHENBERICHT

Internationales Aktenzeichen

PCT/EP 02/14303

A. KLASIFIZIERUNG DES ANMELDUNGSGEGENSTANDES
IPK 7 C12N15/82 C12N5/14

Nach der internationalen Patentklassifikation (IPK) oder nach der nationalen Klassifikation und der IPK

B. RECHERCHIERTE GEBIETE

Recherchierter Mindestprüfstoff (Klassifikationssystem und Klassifikationssymbole)

IPK 7 C12N

Recherchierte aber nicht zum Mindestprüfstoff gehörende Veröffentlichungen, soweit diese unter die recherchierten Gebiete fallen

Während der internationalen Recherche konsultierte elektronische Datenbank (Name der Datenbank und evtl. verwendete Suchbegriffe)

EPO-Internal, BIOSIS, WPI Data, PAJ, EMBASE

C. ALS WESENTLICH ANGESEHENE UNTERLAGEN

Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
X	WO 01 29241 A (CALGENE LLC) 26. April 2001 (2001-04-26)	1-5, 10, 12-14
Y	siehe das ganze Dokument	7-9, 11
X	WO 01 07572 A (UNIV CALIFORNIA ; US AGRICULTURE (US)) 1. Februar 2001 (2001-02-01) siehe insbes. Seiten 12-16	7, 9, 11-14
Y	WO 00 11155 A (UNIV LELAND STANFORD JUNIOR) 2. März 2000 (2000-03-02) siehe S. 5, letzter Abschnitt, S. 30-36, Ansprüche	7-9, 11
	-/--	

☒ Weitere Veröffentlichungen sind der Fortsetzung von Feld C zu entnehmen☒ Siehe Anhang Patentfamilie

* Besondere Kategorien von angegebenen Veröffentlichungen :

A Veröffentlichung, die den allgemeinen Stand der Technik definiert, aber nicht als besonders bedeutsam anzusehen ist

E Älteres Dokument, das jedoch erst am oder nach dem internationalen Anmeldedatum veröffentlicht worden ist

L Veröffentlichung, die geeignet ist, einen Prioritätsanspruch zweifelhaft erscheinen zu lassen, oder durch die das Veröffentlichungsdatum einer anderen im Recherchenbericht genannten Veröffentlichung belegt werden soll oder die aus einem anderen besonderen Grund angegeben ist (wie ausgeführt)

O Veröffentlichung, die sich auf eine mündliche Offenbarung, eine Benutzung, eine Ausstellung oder andere Maßnahmen bezieht

P Veröffentlichung, die vor dem internationalen Anmeldedatum, aber nach dem beanspruchten Prioritätsdatum veröffentlicht worden ist

T Spätere Veröffentlichung, die nach dem internationalen Anmeldedatum oder dem Prioritätsdatum veröffentlicht worden ist und mit der Anmeldung nicht kollidiert, sondern nur zum Verständnis des der Erfindung zugrundeliegenden Prinzips oder der ihr zugrundeliegenden Theorie angegeben ist

X Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann allein aufgrund dieser Veröffentlichung nicht als neu oder auf erfinderischer Tätigkeit beruhend betrachtet werden

Y Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann nicht als auf erfinderischer Tätigkeit beruhend betrachtet werden, wenn die Veröffentlichung mit einer oder mehreren anderen Veröffentlichungen dieser Kategorie in Verbindung gebracht wird und diese Verbindung für einen Fachmann naheliegend ist

G Veröffentlichung, die Mitglied derselben Patentfamilie ist

Datum des Abschlusses der internationalen Recherche

31. März 2003

Absendedatum des internationalen Recherchenberichts

11/04/2003

Name und Postanschrift der internationalen Recherchenbehörde

Europäisches Patentamt, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Bevollmächtigter Beauftragter

Grosskopf, R

C.(Fortsetzung) ALS WESENTLICH ANGESEHENE UNTERLAGEN

Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
X	<p>CORNEILLE SYLVIE ET AL: "Efficient elimination of selectable marker genes from the plastid genome by the CRE-lox site-specific recombination system." PLANT JOURNAL, Bd. 27, Nr. 2, Juli 2001 (2001-07), Seiten 171-178, XP002236641 ISSN: 0960-7412 siehe S. 172, linke Kolonne</p> <p>---</p>	7-10
A	<p>BOGORAD L: "Engineering chloroplasts: an alternative site for foreign genes, proteins, reactions and products" TRENDS IN BIOTECHNOLOGY, ELSEVIER, AMSTERDAM, NL, Bd. 18, Nr. 6, Juni 2000 (2000-06), Seiten 257-263, XP004203651 ISSN: 0167-7799</p> <p>-----</p>	

INTERNATIONALER RECHERCHENBERICHT

Internationales Aktenzeichen
PCT/EP 02/14303

Feld I Bemerkungen zu den Ansprüchen, die sich als nicht recherchierbar erwiesen haben (Fortsetzung von Punkt 2 auf Blatt 1)

Gemäß Artikel 17(2)a) wurde aus folgenden Gründen für bestimmte Ansprüche kein Recherchenbericht erstellt:

1. ☐ Ansprüche Nr.
weil sie sich auf Gegenstände beziehen, zu deren Recherche die Behörde nicht verpflichtet ist, nämlich
2. ☐ Ansprüche Nr.
weil sie sich auf Teile der internationalen Anmeldung beziehen, die den vorgeschriebenen Anforderungen so wenig entsprechen, daß eine sinnvolle internationale Recherche nicht durchgeführt werden kann, nämlich
3. ☐ Ansprüche Nr.
weil es sich dabei um abhängige Ansprüche handelt, die nicht entsprechend Satz 2 und 3 der Regel 6.4 a) abgefaßt sind.

Feld II Bemerkungen bei mangelnder Einheitlichkeit der Erfindung (Fortsetzung von Punkt 3 auf Blatt 1)

Die internationale Recherchenbehörde hat festgestellt, daß diese internationale Anmeldung mehrere Erfindungen enthält:

siehe Zusatzblatt

1. ☐ Da der Anmelder alle erforderlichen zusätzlichen Recherchegebühren rechtzeitig entrichtet hat, erstreckt sich dieser internationale Recherchenbericht auf alle recherchierbaren Ansprüche.
2. ☒ Da für alle recherchierbaren Ansprüche die Recherche ohne einen Arbeitsaufwand durchgeführt werden konnte, der eine zusätzliche Recherchegebühr gerechtfertigt hätte, hat die Behörde nicht zur Zahlung einer solchen Gebühr aufgefordert.
3. ☐ Da der Anmelder nur einige der erforderlichen zusätzlichen Recherchegebühren rechtzeitig entrichtet hat, erstreckt sich dieser internationale Recherchenbericht nur auf die Ansprüche, für die Gebühren entrichtet worden sind, nämlich auf die Ansprüche Nr.
4. ☐ Der Anmelder hat die erforderlichen zusätzlichen Recherchegebühren nicht rechtzeitig entrichtet. Der internationale Recherchenbericht beschränkt sich daher auf die in den Ansprüchen zuerst erwähnte Erfindung; diese ist in folgenden Ansprüchen erfaßt:

Bemerkungen hinsichtlich eines Widerspruchs

- ☐ Die zusätzlichen Gebühren wurden vom Anmelder unter Widerspruch gezahlt.
- ☐ Die Zahlung zusätzlicher Recherchegebühren erfolgte ohne Widerspruch.

WEITERE ANGABEN**PCT/ISA/ 210**

Die internationale Recherchenbehörde hat festgestellt, daß diese internationale Anmeldung mehrere (Gruppen von) Erfindungen enthält, nämlich:

1. Ansprüche: 1-6, 10 (komplett), 11-14 (teilweise)

Verfahren zur sequenzspezifischen Integration einer DNA Sequenz in die plastidäre DNA eines pflanzlichen Genoms unter Verwendung einer Rekombinase, sowie die damit erhaltenen Pflanzen und deren Verwendung

2. Ansprüche: 7 (komplett), 9, 11-14 (teilweise)

Expressionskassette enthaltend eine Nukleinsäuresequenz kodierend für eine Resolvase/Invertase unter der Kontrolle eines in pflanzlichen Plastiden aktiven Promotors, pflanzliche Organismen, die diese Kassette enthalten und deren Verwendung

3. Ansprüche: 8 (komplett), 9, 11-14 (teilweise)

Expressionskassette enthaltend eine Nukleinsäuresequenz kodierend für eine Fusionsprotein aus einer Plastidenlokalisationssequenz und eine Resolvase/Invertase unter der Kontrolle eines im pflanzlichen Zellkern aktiven Promotors, pflanzliche Organismen, die diese Kassette enthalten und deren Verwendung

INTERNATIONALER RECHERCHENBERICHT

Angaben zu Veröffentlichungen, die zur selben Patentfamilie gehören

Internationales Aktenzeichen

PCT/EP 02/14303

Im Recherchenbericht angeführtes Patentdokument	Datum der Veröffentlichung	Mitglied(er) der Patentfamilie	Datum der Veröffentlichung
WO 0129241 A	26-04-2001	AU 1090701 A	30-04-2001
		BR 0014858 A	16-07-2002
		EP 1222296 A2	17-07-2002
		WO 0129241 A2	26-04-2001
WO 0107572 A	01-02-2001	AU 6364800 A	13-02-2001
		EP 1234022 A2	28-08-2002
		JP 2003505065 T	12-02-2003
		WO 0107572 A2	01-02-2001
WO 0011155 A	02-03-2000	AU 5898599 A	14-03-2000
		WO 0011155 A1	02-03-2000